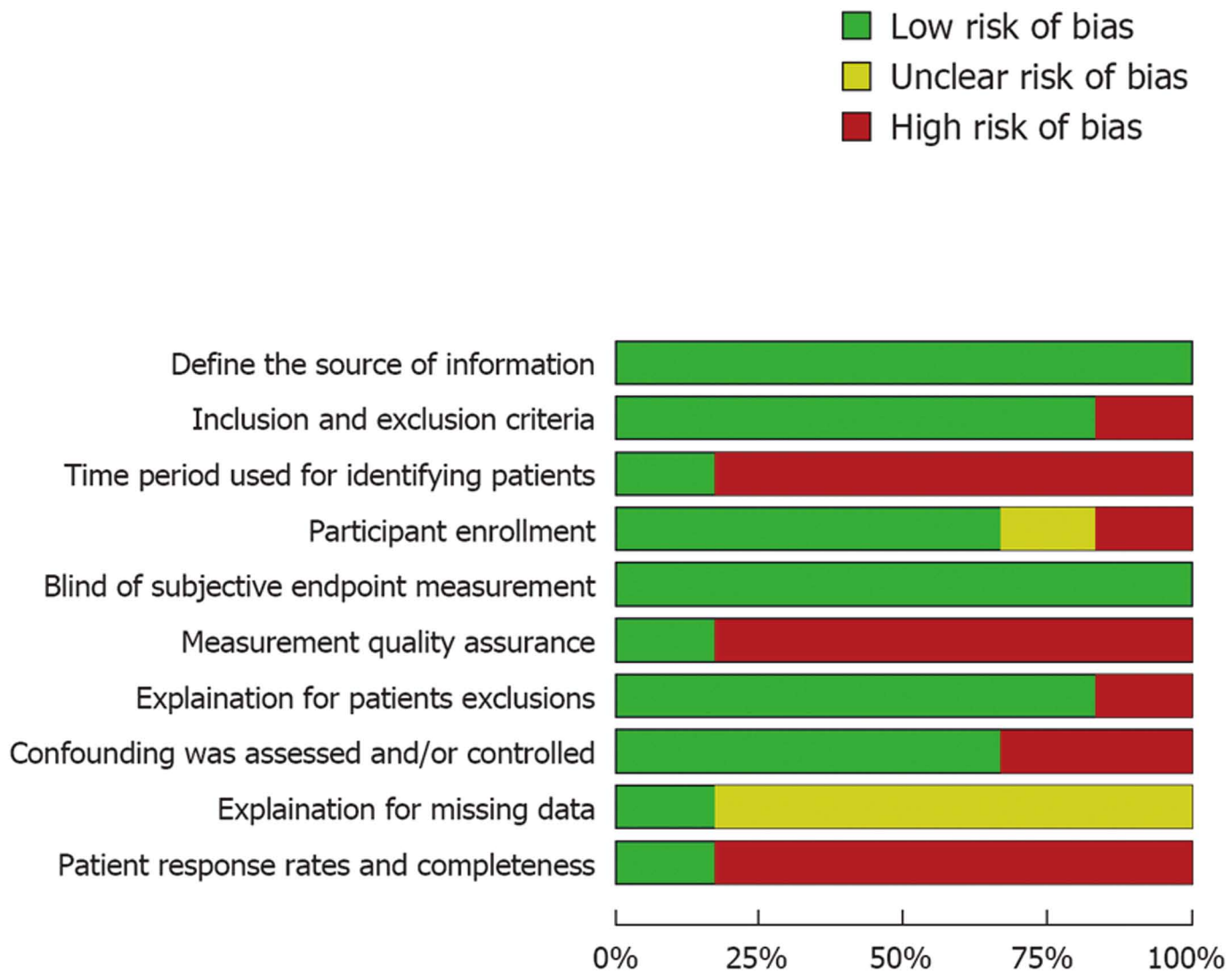
ISSN 2308-3840 (online)

# World Journal of *Meta-Analysis*

*World J Meta-Anal* 2013 August 26; 1(2): 57-96




Baishideng Publishing Group
www.wjgnet.com

WJMA World Journal of Meta-Analysis

A peer-reviewed, online, open-access journal of meta-analysis

# Editorial Board

2013-2018

The *World Journal of Meta-Analysis* Editorial Board consists of 402 members, representing a team of worldwide experts in clinical meta-analysis research. They are from 41 countries, including Argentina (2), Australia (3), Austria (1), Belgium (5), Brazil (11), Canada (17), Chile (3), China (121), Croatia (1), Egypt (1), Finland (4), France (2), Germany (9), Greece (11), Hungary (2), India (12), Iran (2), Ireland (1), Israel (2), Italy (41), Japan (5), Lithuania (1), Malaysia (1), Netherlands (8), New Zealand (1), Norway (1), Peru (1), Poland (4), Portugal (6), Romania (1), Saudi Arabia (4), Singapore (3), South Africa (1), South Korea (7), Spain (8), Sri Lanka (2), Sudan (1), Switzerland (2), Thailand (4), Turkey (3), United Kingdom (22), and United States (65).

Zhi-Wei He, *Dongguan*
Ben He, *Shanghai*
Guo-Wei He, *Tianjin*
G Huang, *Shanghai*
Bing-Yang Ji, *Beijing*
Jing Jiang, *Changchun*
Joey Sum Wing Kwong, *Hong Kong*
Wei-Dong Leng, *Shiyan*
Le-Qun Li, *Nanning*
Xiao-Ping Li, *Chengdu*
Jian-Sheng Li, *Zhengzhou*
Jing-Cheng Li, *Chongqing*
Jun-Sheng Li, *Nanjing*
Yan-Yan Li, *Nanjing*
Hua Liu, *Nanchong*
Tong Liu, *Tianjin*
Ai-Ping Lu, *Hong Kong*
Ying Luo, *Kunming*
Chao Ma, *Shanghai*
Jie Ma, *Xi'an*
Xiang-Yu Ma, *Chongqing*
Yan-Lei Ma, *Shanghai*
Wei Nie, *Shanghai*
Wen-Quan Niu, *Shanghai*
Wen-Sheng Pan, *Hangzhou*
Shi-Qiang Shen, *Wuhan*
Xiang-Chun Shen, *Guiyang*
Rui-Hua Shi, *Nanjing*
Yong-Bing Shi, *Suzhou*
Ke-Qing Shi, *Wenzhou*
Marcelo Signorini, *Hong Kong*
Zhi-Yuan Song, *Chongqing*
Qing-Min Sun, *Nanjing*
Yi-Hong Sun, *Beijing*
Shi-Qiao Tan, *Chengdu*
SQ Tan, *Chengdu*
Yong Tang, *Tianjin*
Jiu-Lai Tang, *Hefei*
Na-Ping Tang, *Shanghai*
Jian-Cheng Tu, *Wuhan*
Bin Wang, *Beijing*
Dao-Rong Wang, *Yangzhou*
Jing Wang, *Changshu*
Yu-Ting Wang, *Chengdu*
Xi-Shan Wang, *Harbin*
Fu-Zhou Wang, *Nanjing*
Bing Xia, *Wuhan*
Hong-Xia Wang, *Shanghai*
Zhen-Ning Wang, *Shenyang*
Na Wang, *Shijiazhuang*
Shu-Kui Wang, *Wuhan*
Xing-Huan Wang, *Wuhan*
Wei Wang, *Wuxi*
Cong-Xia Wang, *Xi'an*
Feng Xie, *Shanghai*
Zi-Qiang Xin, *Beijing*
Dan Xing, *Tianjin*
Jun Xiong, *Nanchang*
Xi-Ping Xu, *Guangzhou*
Lin Xu, *Nanjing*
Zhuo-Qun Xu, *Wuxi*
Hui-Ping Xue, *Shanghai*
Tian Yang, *Changchun*
Shuan-Ying Yang, *Xi'an*
Yi-Cong Ye, *Beijing*
Yan-Wei Yin, *Beijing*
Zi Yin, *Guangzhou*
Ym Yin, *Nanjing*
Bin Yu, *Guangzhou*
Yun-Xian Yu, *Hangzhou*
Ling Zhang, *Beijing*
Bei-Bei Zhang, *Chengdu*
Li-Li Zhang, *Chongqing*
Qiu Zhang, *Hefei*
Shuo Zhang, *Shenyang*
You-Cheng Zhang, *Lanzhou*
Jian Zhang, *Shanghai*
Jun-Hua Zhang, *Tianjin*
Yu-Rong Zhang, *Xi'an*
Zhong-Heng Zhang, *Jinhua*
Hai-Tao Zhao, *Beijing*
Pan Zhao, *Beijing*
Yu-Lan Zhao, *Shanghai*
Jie-Jiao Zheng, *Shanghai*
Xue-Sheng Zheng, *Shanghai*
Guo-Qing Zheng, *Wenzhou*
Ming-Hua Zheng, *Wenzhou*
Cui-Hong Zheng, *Wuhan*
Lai-Ping Zhong, *Shanghai*
Tian-Biao Zhou, *Nanning*
Peng Zhou, *Shanghai*
Ping Zhou, *Wuhan*
Kun-Ju Zhu, *Zhanjiang*
Xiao-Ping Zou, *Nanjing*

**Croatia**

Miljenko Franic, *Zagreb*

**Egypt**

Ashraf Fawzy Nabhan, *Cairo*

**Finland**

Jouni JK Jaakkola, *Oulu*
Ville Kyto, *Turku*
Jouko Miettunen, *Oulu*
R Quansah, *Oulu*

**France**

Alain Braillon, *Amiens*
Francesco Fiorica, *Ferrara*

**Germany**

Tonio Ball, *Freiburg*
Robert Bergholz, *Hamburg*
Jan Brunkwall, *Cologne*
Holger Cramer, *Essen*
Joseph Kambeitz, *Munich*
Sascha Meyer, *Homburg*
Thomas Nickl-Jockschat, *Aachen*
Martin Pinquart, *Marburg*
Robert Schier, *Cologne*

**Greece**

Vangelis G Alexiou, *Athens*
Stefanos Bonovas, *Athens*
Athanasios Papatsoris, *Athens*
Irini Chatziralli, *Athens*
Dimitrios Daoussis, *Patras*
Pagona Lagiou, *Athens*
John Goudakos, *Thessaloniki*
Savas Grigoriadis, *Thessaloniki*
Konstantinos A Toulis, *Thessaloniki*
Sotirios Tsiodras, *Athens*
Nikolaos Tsoukalas, *Athens*

**Hungary**

Balazs Gyorffy, *Budapest*
Istvan Wittmann, *Pecs*

**India**

Ritesh Agarwal, *Chandigarh*
Giridhara R Babu, *Bangalore*
Subho Chakrabarti, *Chandigarh*
Y Madhavi, *New Delhi*
Tanu Midha, *Kanpur*
Kameshwar Prasad, *New Delhi*
Kaushal Kishor Prasad, *Chandigarh*
Krishna Undela, *Mysore*
Singh Rajender, *Lucknow*
Vinod Ravindran, *Kozhikode*
V Shetty, *Mumbai*
R Umaya Suganthi, *Bangalore*

**Iran**

Nejat Mahdieh, *Ilam*
Sadeghi Ramin, *Mashhad*

**Ireland**

Ian Conrick-Martin, *Dublin*

**Israel**

Uri Kopylov, *Ramat Gan*
Meir Lotan, *Kfar-Saba*

**Italy**

Umberto Aguglia, *Catanzaro*
Alessandro Antonelli, *Pisa*
Annalisa Blasetti, *Chieti*
Francesco Brigo, *Merano*
Emanuele Cereda, *Pavia*
Emanuele Cigna, *Rome*

Roberto Cirocchi, *Terni*
Bernardo Cortese, *Milano*
Alessandro Cucchetti, *Bologna*
Gianfranco Damiani, *Rome*
Fabrizio D'Ascenzo, *Turin*
Massimo Del Fabbro, *Milano*
Valeria Fadda, *Florence*
Alessandro Fancellu, *Sassari*
Giuseppe Ferrante, *Milan*
Virginia Festa, *Rome*
Francesco Fiorica, *Ferrara*
Guglielmo Giraldi, *Rome*
Jenny Guidi, *Bologna*
Lorenzo Loffredo, *Rome*
Andrea Messori, *Firenze*
Eliano Pio Navarese, *Bydgoszcz*
Stefano Omboni, *Solbiate Arno*
Alvisa Palese, *Udine*
Stefano Palomba, *Reggio Emilia Stefano*
Carlo Perricone, *Rome*
Mario Petretta, *Naples*
Alessandro Pezzini, *Brescia*
Gianluca Pontone, *Milan*
Paolo Emilio Puddu, *Rome*
Andrea Rognoni, *Novara*
Giuseppe Scalabrino, *Milan*
Città della Salute e della Scienza, *Turin*
Fabrizio Sgolastra, *L'Aquila*
Maria Lucia Specchia, *Rome*
Stefano Trastulli, *Terni*
Fabio Tine, *Palermo*
Nereo Vettoretto, *Chiari*
Alberto Vianello, *Perugia*
Luigi Zorcolo, *Cagliari*

**Japan**

Nguyen Tien Huy, *Nagasaki*
Hiroharu Kamioka, *Tokyo*
Koji Kawakami, *Kyoto*
Keitaro Matsuo, *Nagoya*
Kazushi Okamoto, *Nagoya*

**Lithuania**

Edmundas Kadusevicius, *Kaunas*

**Malaysia**

SP Pani, *Ipoh*

**Netherlands**

Michel van den Bekerom, *Amsterdam*
Dimitra Dodou, *Delft*
D Haverkamp, *Amsterdam*
Vassilios Koussoulas, *Drachten*
BJ Polder, *Nijmegen*
Theo Stijnen, *Leiden*
RNM Weijers, *Amsterdam*
Joost de Winter, *Delft*

**New Zealand**

Shaofeng Li, *Auckland*

**Norway**

Eivind Berge, *Oslo*

**Peru**

Rafael Bolaños Díaz, *Lima*

**Poland**

Maciej Banach, *Lodz*
Krzysztof Jonderko, *Sosnowiec*
Jolanta Lissowska, *Warsaw*
Maciej Plaszewski, *Warsaw*

**Portugal**

Daniel Caldeira, *Lisboa*
J Costa, *Lisbon*
Ana Miguel, *Coimbra*
Manuel Morgado, *Covilh*
Bárbara Peleteiro, *Porto*
Rui Torres, *Paredes*

**Romania**

Fratila Ovidiu, *Oradea*

**Saudi Arabia**

Hazem M Al-Mandeel, *Riyadh*
Ezzeldin M Ibrahim, *Jeddah*
Mutahir A Tunio, *Riyadh*
Hayfaa A Wahabi, *Riyadh*

**Singapore**

Nikos LD Chatzisarantis, *Singapore*
Edwin Choon Wyn Lim, *Singapore*
Roger Ho, *Singapore*

**South Africa**

Alaine Umubyeyi Nyaruhirira, *Pretoria*

**South Korea**

Jung-Hee Kim, *Cheonan*
Hyangsook Lee, *Seoul*
Myeong Soo Lee, *Daejeon*
Chi-Un Pae, *Bucheon*
Yong Hyun Park, *Yangsan*
Jae Hong Seo, *Seoul*
Yong Sang Song, *Seoul*

**Spain**

Pablo Avanzas, *Oviedo*
Joan Cid, *Barcelona*
Joaquin de Haro, *Madrid*
Joan Guardia-Olmos, *Barcelona*
Nabil Halaihlel, *Zaragoza*
JA Monge-Argilés, *Alicante*
Raul Moreno, *Madrid*
Inés Velasco, *Aracena*

**Sri Lanka**

Ranil Jayawardena, *Colombo*
Priyanga Ranasinghe, *Colombo*

**Switzerland**

Jay P Singh, *Zurich*
Giorgio Treglia, *Bellinzona*

**Sudan**

Samir MH Shaheen, *Khartoum*

**Thailand**

Chuenjid Kongkaew, *Phitsanulok*
Manop Pithukpakorn, *Bangkok*
Piyamitr Sritara, *Bangkok*
Surasak Saokaew, *Phayao*

**Turkey**

Nese Demirturk, *Afyonkarahisar*
Nilüfer Ozabaci, *Eskisehir*
Ilke Sipahi, *Istanbul*

**United Kingdom**

Omar M Aboumarzouk, *Wales*
Abeer Al-Namankany, *London*
Ernest A Azzopardi, *Cardif*
Umberto Benedetto, *Cambridge*
Joanne Brooke, *London*
Noriko Cable, *London*
David Chan, *Cardiff*
YC Cheong, *Southampton*
Andrew Currie, *Harrow*
G Nabi, *Dundee*

WJMA

World Journal of Meta-Analysis

Contents

Quarterly Volume 1 Number 2 August 26, 2013

**META-ANALYSIS**

*World Journal of Meta-Analysis*
**Volume 1 Number 2 August 26, 2013**

## Contents





NAME OF JOURNAL
*World Journal of Meta-Analysis*

ISSN
ISSN 2308-3840 (online)

LAUNCH DATE
May 26, 2013

FREQUENCY
Quarterly



EDITORIAL OFFICE
Jin-Lei Wang, Director
Xiu-Xia Song, Vice Director
*World Journal of Meta-Analysis*
Room 903, Building D, Ocean International Center,
No. 62 Dongsihuan Zhonglu, Chaoyang District,
Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: wjma@wjgnet.com
http://www.wjgnet.com

PUBLISHER
Baishideng Publishing Group Co., Limited
Flat C, 23/F., Lucky Plaza,
315-321 Lockhart Road, Wan Chai,
Hong Kong, China
Fax: +852-6555-7188
Telephone: +852-3177-9906
E-mail: bpgoffice@wjgnet.com
http://www.wjgnet.com

PUBLICATION DATE
August 26, 2013



INSTRUCTIONS TO AUTHORS
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm

ONLINE SUBMISSION
http://www.wjgnet.com/esps/

World Journal of Meta-Analysis

Online Submissions: http://www.wjgnet.com/esps/
wjma@wjgnet.com
doi:10.13105/wjma.v1.i2.57

*World J Meta-Anal* 2013 August 26; 1(2): 57-77
ISSN 2308-3840 (online)


META-ANALYSIS

# Dose-response relationship of lung cancer to amount smoked, duration and age starting

John S Fry, Peter N Lee, Barbara A Forey, Katharine J Coombs

John S Fry, Peter N Lee, Barbara A Forey, Katharine J Coombs, PN Lee Statistics and Computing Ltd., Sutton, Surrey SM2 5DA, United Kingdom
Author contributions: Lee PN, Fry JS and Forey BA planned the study; Literature searches were carried out by Coombs KJ, assisted by Lee PN and Forey BA; Data entry was carried out by Coombs KJ and checked by Forey BA, or carried out by Forey BA and checked by Lee PN; Where appropriate, difficulties in interpreting published data or in the appropriate methods for derivation of RRs were discussed by Forey BA and Lee PN; The statistical analyses were conducted by Fry JS along lines discussed and agreed with Lee PN; Lee PN and Fry JS jointly drafted the paper, which was critically reviewed by Forey BA and Coombs KJ.
Supported by Philip Morris Products SA
Correspondence to: Peter N Lee, MA, Director, PN Lee Statistics and Computing Ltd., 17 Cedar Road, Sutton, Surrey SM2 5DA, United Kingdom. peterlee@pnlee.co.uk
Telephone: +44-20-86428265 Fax: +44-20-86422135
Received: April 2, 2013 Revised: May 9, 2013
Accepted: August 4, 2013
Published online: August 26, 2013

## Abstract

**AIM:** To quantify smoking/lung cancer relationships accurately using parametric modelling.

**METHODS:** Using the International Epidemiological Studies on Smoking and Lung Cancer database of all epidemiological studies of 100+ lung cancer cases published before 2000, we analyzed 97 blocks of data for amount smoked, 35 for duration of smoking, and 27 for age started. Pseudo-numbers of cases and controls (or at risk) estimated from RRs by dose level formed the data modelled. We fitted various models relating $\log_e$ RR to dose (d), including $\beta d$, $\beta d^{\gamma}$ and $\beta \log_e (1 + \omega d)$, and investigated goodness-of-fit and heterogeneity between studies.

**RESULTS:** The best-fitting models for $\log_e$ RR were $0.833 \log_e [1 + (8.1c/10)]$ for cigarettes/d (c), $0.792 (y/10)^{0.74}$ for years smoked (y) and $0.176 [(70 - a)/10]^{1.44}$ for age of start (a). Each model fitted well overall, though some blocks misfitted. RRs rose from 3.86 to 22.31 between c = 10 and 50, from 2.21 to 13.54 between y = 10 and 50, and from 3.66 to 8.94 between a = 30 and 12.5. Heterogeneity ($P < 0.001$) existed by continent for amount, RRs for 50 cigarettes/d being 7.23 (Asia), 26.36 (North America) and 22.16 (Europe). Little heterogeneity was seen for duration of smoking or age started.

**CONCLUSION:** The models describe the dose-relationships well, though may be biased by factors including misclassification of smoking status and dose.



**Key words:** Smoking; Lung neoplasms; Dose-response; Meta-analysis; Review; Amount smoked; Duration of smoking; Age at starting to smoke

**Core tip:** This paper, for the first time, meta-analyses smoking/lung cancer dose-relationships. Based on data from 71 studies published before 2000, single parameter models were fitted to summarize how the RR increased with increasing amount smoked, longer duration of smoking, and earlier age of starting to smoke. Overall, the models fitted well. Little heterogeneity was seen for duration of smoking or age of start, but the rise in RR with amount smoked was much steeper in North America and Europe than in Asia. The fitted models can be used to more precisely estimate the lung cancer risk from smoking.

Fry JS, Lee PN, Forey BA, Coombs KJ. Dose-response relationship of lung cancer to amount smoked, duration and age starting. *World J Meta-Anal* 2013; 1(2): 57-77 Available from: URL: http://www.wjgnet.com/2308-3840/full/v1/i2/57.htm DOI: http://dx.doi.org/10.13105/wjma.v1.i2.57

## INTRODUCTION

We recently carried out a systematic review[1] of the evidence relating smoking to lung cancer incorporating all 287 studies published before 2000 involving a minimum of 100 lung cancer cases. We refer to this as "our earlier review". In that review, we assessed evidence concerning amount smoked per day, duration of smoking, and age of starting to smoke. Data are typically available as blocks of RRs for differing levels of the dose-response measure, each compared to never smokers. Comparing meta-analysis estimates for low, medium and high exposure, we clearly demonstrated a dose-response existed. For example, for amount smoked by current smokers, random-effects RR estimates are 4.71 (95%CI: 4.14-5.37, $n$ = 86) for about 5 cigs/d, 9.83 (95%CI: 8.60-11.24, $n$ = 54) for about 20 cigs/d, and 17.10 (95%CI: 14.62-19.99, $n$ = 62) for about 45 cigs/d. Here "about 5 cigs/d" combined results for dose ranges including 5 but not 20 cigs/d, "about 20 cigs/d" considered ranges including 20 but not 5 or 45 cigs/d, and "about 45 cigs/d" ranges including 45 but not 20 cigs/d. This approach has limitations. First, formal statistical comparison of the RRs at the different levels is not possible as the RRs are not independent, having the same denominator. Second, the analyses do not use all the information available. Thus, results for ranges wholly between 5 and 20 cigs/d or wholly between 20 and 45 cigs/d are ignored, as are results for ranges covering two or more of the "key values" of 5, 20 and 45 cigs/d. Also, linearity, or other shapes of the relationship, is not assessed. Dose-response relationships for years quit are considered in a separate paper[2].

Here, we study dose-response in more detail by fitting models to the various dose-response blocks to estimate parameters which can be meta-analyzed and used to assess heterogeneity. We follow the approach previously used[3] to quantify the dose-response relationship between environmental tobacco smoke exposure and lung cancer risk, developing a variant of it for age of starting. We restrict attention to the data considered in our earlier review[1]. Rather than also considering results for ever smokers, we restrict attention to current smokers, giving a more homogeneous dataset and one showing a stronger dose-relationship. All our analyses are of overall lung cancer risk, no attempt being made in the present paper to fit models for specific histological types.

## MATERIALS AND METHODS

### The International Evidence on Smoking and Lung Cancer database

All analyses use the International Evidence on Smoking and Lung Cancer database, fully described in our earlier review[1]. Papers considered were published before 2000, described studies of 100+ cases, and provided RR estimates for one or more smoking indices. We use the term RR generically to describe alternative RR estimates, *e.g.*, odds ratio or hazard ratio. Lee *et al*[1] gives details of the structure and data entry rules for the database.

### Data selection and blocks

The data considered here comprise blocks of RRs, each relative to never smokers, for all lung cancer (or occasionally near equivalent definitions, each including squamous cell carcinoma and adenocarcinoma) for three measures of dose-response among current smokers: amount smoked, duration and age of starting. Where possible, blocks by sex or by sex and race were considered. Except for amount smoked, blocks by age were considered, if available. Covariate-adjusted RRs were preferred to unadjusted RRs. Each block includes an estimate of the RR and 95%CI: for each level of the measure. The data recorded per block included study type, sex, location, publication year, age range (at baseline for prospective studies), product smoked [any product, cigarettes +/- other products (*i.e.*, pipes, cigars), cigarettes only], never smoker definition (never any product, never cigarettes). For each RR, the range of the measure was also recorded.

### Pseudo-numbers

We used the method of Hamling *et al*[4] on each block to estimate the pseudo-table of numbers of cases, and either controls (for case-control studies) or at risk (for prospective studies) which correspond to the observed RRs and 95%CIs. The method was applied even to unadjusted RRs. This estimation requires, in addition to the given RRs and 95%CIs, estimates of the proportion of never smokers among the controls/at risk and of the ratio of total controls/at risk to total cases, as well as starting values for the numbers of never smoking cases and controls/at risk. These estimates were also recorded on the database. The pseudo-table forms the basic data for fitting the models used, and estimating the overall current smoking RR.

### Midpoints for levels of exposure

For amount smoked, midpoint estimates for each exposure level were derived using standard distributions, as described by Fry *et al*[3] when relating lung cancer risk to amount smoked by the husband. For US studies, the distribution derived from published data for two large CPS I and CPS II studies[5], while for non-US studies, it was that given in Appendix III of International Smoking Statistics[6,7].

For duration of smoking and age of starting the midpoints were based on US NHANES III[8], selecting data for subjects for the given sex, age and range of values of the relevant dose-response measure.

### Statistical models

For each measure, the data analyzed consist of blocks, each containing the pseudo-numbers and the estimated midpoint exposures for each of $\ell$ exposure levels, and for never smokers. The methodology varies by dose-response measure, as described below.

**Table 1 Models used to relate risk to dose**

| | | | |
|---|---|---|---|
| $\log_e$ (RR) | = | $\beta_1 d$ | (linear) |
| $\log_e$ (RR) | = | $\beta_1 d + \beta_2 d^2$ | (quadratic) |
| $\log_e$ (RR) | = | $\beta_1 d + \beta_2 d^2 + \beta_3 d^3$ | (cubic) |
| $\log_e$ (RR) | = | $\beta_1 d^Y$ | (power) |
| $\log_e$ (RR) | = | $\beta_1 \log_e$ (d) | (log) |
| $\log_e$ (RR) | = | $\beta_1 \exp_e$ (d) | (exponential) |
| $\log_e$ (RR) | = | $\beta_1 \log_e$ (1 + Wd) | (log-with-baseline) |

### Amount smoked

The Greenland and Longnecker method[9,10] was used to fit functional forms relating RR to dose (midpoint amount smoked). We fitted the models expressing dose, d, in units of 10 cigs/d.

In the simplest application, the RR is predicted by $\log_e$ RR (d) = βd, β and SE (β) are estimated separately per block, and estimates of β and SE (β) are then combined using inverse-variance weighted random-effects or fixed-effects meta-analysis[11]. This model implies that a fixed dose increment increases risk by a fixed factor. The method can be used with d replaced by a function of d, such as $d^{1/2}$, $d^2$, or log (d + 1).

Greenland *et al*[9] describe a more general, "pool-first", method in which all the blocks are considered in a single analysis. The method gives the same results for the model log RR = βd, but allows direct fitting for other functional forms.

As the best model was initially unclear, we first tried various models (Table 1) using the pool-first method, comparing deviances to assess which models fitted the overall data better. For the "power" and "log-with-baseline" models, the parameters Y or W could not be fitted directly, but an iterative method was adopted, comparing deviances for a range of values.

For the models with lowest deviance, the simpler approach was then used to estimate $\beta_1$ (and $\beta_2$, $\beta_3$) and its standard error (SE) for each block. For a particular model, goodness-of-fit for a block was tested by comparing observed and fitted number of cases (and for case-control studies also the observed number of controls) at each level of amount smoked (including never smokers). The fitted values were estimated as described in Goodness of fit[11]. As also described there, the sum of (observed - fitted)$^2$/fitted over levels was taken as an approximate chi-squared on ℓ - 1 degrees of freedom (df) for prospective studies or on 2ℓ - 1 df for case-control studies. Information on overall goodness-of-fit was derived by summing observed and fitted values over blocks for never smokers and for specified levels of amount smoked, and similarly deriving an approximate chi-squared statistic. Plots of observed and predicted RRs per block were also examined.

### Duration of smoking

The approach used was as for amount smoked.

### Age of starting to smoke

For smokers of a given age, age of starting (a) and duration (y) are directly related. We used the same basic approach, replacing y by 70 - a to produce a duration-like measure. As this produced a relatively good fit, we did not attempt sensitivity analyses replacing y by 60 - a or 80 - a.

### Regression analyses

Sources of heterogeneity were studied by inverse-variance weighted regression of β. Between block variation was examined one factor at a time (simple regression), and using forward stepwise methods. The factors used were study type, sex, location, publication year, midpoint age (at baseline for prospective studies), smoking product and study size. The deviance of the fitted models indicated the extent to which heterogeneity was explained.

### Statistical analysis

No multiple testing adjustments were made, significant being defined as $P < 0.05$. However, results showing stronger evidence of a relationship ($P < 0.01$, or $P < 0.001$), and sometimes weaker evidence ($P < 0.1$) are also distinguished, where appropriate. All data entry and most statistical analyses were carried out using ROELEE version 3.1 (available from PN Lee Statistics and Computing Ltd., 17 Cedar Road, Sutton, Surrey SM2 5DA, United Kingdom). Some analyses used Excel 2003.

## RESULTS

### Studies considered

For each of the 71 studies providing the data used, Studies[11], gives the six character reference code (REF); a brief description incorporating the location, characteristics of the population studied, study design, and study duration; the total number of lung cancers studied; and the measures for which data are analyzed.

### Amount smoked

Details of blocks used for each measure are given in Blocks[11]. These include study REF, sex (and where applicable race), study type, location, product smoked, definition of unexposed group, adjustment factors used, current smoker RR, and total numbers of cases in smokers.

The 97 blocks derive from 69 studies, 45 providing results for a single block, 22 results by sex, and 2 (DORGAN and HUMBLE) results by sex and race. 55 blocks (56.7%) are for males, 34 (35.1%) females, and 8 (8.2%) both sexes. 48 (49.5%) are from prospective studies. 43 (44.3%) are from North American studies, 32 (33.0%) from Europe and 17 (17.5%) from Asia, the remaining 5 (5.2%) from South America, Africa or Australasia. Five different combinations of product *vs* unexposed occur: cigarettes ± other products *vs* never any product (32 blocks 33.0%), cigarettes ± other products *vs* never cigarettes (29, 29.9%), any product *vs* never any product (20, 20.6%), cigarettes only *vs* never any product (13, 13.4%) and cigarettes only *vs* never cigarettes for (3, 3.1%). Of the 8 blocks for sexes combined, 4 (50.0%) concern RRs adjusted for sex, while 64 of the full 97 blocks (66.0%)

**Table 2 Amount smoked by current smokers (cigarettes per day) - dose-response data**

| Block: Study | Amount smoked groupings[1] | Mean values | RRs[2] |
|---|---|---|---|
| 1: AKIBA | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 3.50, 6.10, 19.10 M |
| 2: AKIBA | 1-14, 15+ | 8.11, 24.09 | 3.60, 5.80 M |
| 3: ARCHER | 1-19, 20, 21+ | 10, 20, 31.83 | 3.53, 6.09, 8.52 M |
| 4: AXELSS | 20 | 20 | 43.30 |
| 5: BENSHL | 1-9, 10-19, 20+ | 4.85, 12.73, 26.03 | 4.00, 9.05, 10.95 M |
| 6: BEST | 1-9, 10-20, 21+ | 4.85, 15.92, 31.83 | 10.00, 16.41, 17.31 M |
| 7: BOUCOT | 1-20, 21-40, 41+ | 13.38, 29.02, 53.33 | 54.09, 78.56, 161.70 M |
| 8: BRETT | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 2.55, 4.25, 8.00 M |
| 9: BROSS | 1-20, 21+ | 13.38, 31.83 | 4.91, 7.20 M |
| 10: BUFFLE | 1-19, 20, 21+ | 10.20, 31.83 | 5.60, 11.84, 22.10 M |
| 11: CEDERL | 1-7, 8-15, 16+ | 4.29, 11.61, 25.41 | 3.40, 7.50, 11.90 M |
| 12: CEDERL | 1-7, 8-15, 16+ | 4.29, 11.61, 25.41 | 2.83, 7.74, 7.56 |
| 13: CHANG | 1-10, 11-20, 21+ | 7.08, 17.67, 31.83 | 5.02, 10.60, 8.26 |
| 14: CHANG | 1-10, 11-20, 21+ | 7.08, 17.67, 31.83 | 3.03, 4.87, 8.21 M |
| 15: CHOW | 1-19, 20-29, 30+ | 10, 21.35, 38.39 | 13.88, 21.87, 44.48 M |
| 16: COMSTO | 1-19, 20-39, 40+ | 10, 22.90, 45.71 | 12.42, 18.16, 24.92 M |
| 17: COMSTO | 1-19, 20-39, 40+ | 10, 22.90, 45.71 | 7.45, 17.35, 13.27 |
| 18: CORREA | 1-20, 21+ | 13.38, 31.83 | 9.30, 25.30 M |
| 19: CPSI | 1-9, 10-19, 20-39, 40+ | 4.85, 12.73, 22.90, 45.71 | 4.51, 8.41, 14.30, 17.49 M |
| 20: CPSII | 1-9, 10-19, 20, 21-39, 40, 41+ | 4.85, 12.73, 20, 26.71, 40, 53.33 | 12.22, 14.52, 21.59, 22.72, 24.14, 45.52 M |
| 21: CPSII | 1-9, 10-19, 20, 21-39, 40, 41+ | 4.85, 12.73, 20, 26.71, 40, 53.33 | 3.89, 8.33, 14.21, 21.40, 19.31, 18.22 |
| 22: DARBY | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 73.47, 95.43, 142.69 M |
| 23: DARBY | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 15.70, 21.50, 41.62 M |
| 24: DEAN3 | 1-12, 13-22, 23+ | 7.65, 18.42, 33.04 | 5.46, 7.42, 21.66 M |
| 25: DEAN3 | 1-12, 13-22, 23+ | 7.65, 18.42, 33.04 | 3.16, 8.42, 24.24 M |
| 26: DEKLER | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 19.40, 23.00, 32.50 M |
| 27: DOLL2 | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 5.20, 10.60, 22.40 M |
| 28: DOLL2 | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 1.29, 6.43, 29.71 M |
| 29: DORANT | 1-9, 10-19, 20+ | 4.85, 12.73, 26.03 | 8.52, 27.22, 36.24 M |
| 30: DORGAN | 1-19, 20+ | 10, 26.03 | 9.13, 20.65 M |
| 31: DORGAN | 1-19, 20+ | 10, 26.03 | 26.67, 72.46 M |
| 32: DORGAN | 1-19, 20+ | 10, 26.03 | 6.55, 24.13 M |
| 33: DORGAN | 1-19, 20+ | 10, 26.03 | 7.43, 41.43 M |
| 34: DORN | 1-9, 10-20, 21-39, 40+ | 4.85, 15.92, 26.71, 45.71 | 4.02, 9.92, 17.19, 22.75 M |
| 35: ENGELA | 1-4, 5-9, 10-14, 15-19, 20+ | 2.5, 6.5, 10.88, 15.83, 26.03 | 1.40, 4.10, 7.00, 11.00, 15.00 M |
| 36: ENGELA | 1-4, 5-9, 10-14, 15+ | 2.5, 6.5, 10.88, 24.09 | 12.00, 12.00, 24.00, 26.00 |
| 37: ENSTRO | 1-9, 10-19, 20, 21-39, 40+ | 4.85, 12.73, 20, 26.71, 45.71 | 4.74, 7.68, 13.65, 16.08, 19.41 M |
| 38: ENSTRO | 1-9, 10-19, 20, 21+ | 4.85, 12.73, 20, 31.83 | 2.15, 4.31, 9.48, 16.47 M |
| 39: GAO2 | 1-19, 20-29, 30+ | 10, 21.35, 38.39 | 3.36, 7.54, 10.63 M |
| 40: GILLIS | 1-14, 15-24, 25-34, 35-49, 50+ | 8.11, 19.19, 28.13, 39, 53.33 | 4.50, 7.60, 8.60, 9.70, 7.80 |
| 41: HAENSZ | 1-20, 21+ | 13.38, 31.83 | 1.77, 5.15 M |
| 42: HAMMO2 | 1-19, 20+ | 10.00, 26.03 | 9.15, 10.39 M |
| 43: HAMMON | 1-9, 10-20, 21-39, 40+ | 4.85, 15.92, 26.71, 45.71 | 7.44, 8.42, 17.91, 20.64 M |
| 44: HIRAYA | 1-9, 10-19, 20+ | 4.85, 12.73, 26.03 | 2.06, 4.00, 6.24 M |
| 45: HIRAYA | 1-9, 10-19, 20+ | 4.85, 12.73, 26.03 | 2.25, 2.56, 4.47 M |
| 46: HITOSU | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 2.08, 2.82, 4.68 M |
| 47: HITOSU | 1-14, 15+ | 8.11, 24.09 | 3.11, 3.17 M |
| 48: HOLE | 1-14, 15-24, 25-34, 35+ | 8.11, 19.19, 28.13, 44.38 | 5.47, 8.90, 10.75, 7.49 |
| 49: HUMBLE | 1-19, 20+ | 10, 26.03 | 9.20, 24.70 M |
| 50: HUMBLE | 1-19, 20+ | 10, 26.03 | 11.60, 26.10 M |
| 51: HUMBLE | 1-19, 20+ | 10, 26.03 | 19.20, 16.00 |
| 52: HUMBLE | 1-19, 20+ | 10, 26.03 | 18.50, 36.90 M |
| 53: KAISE2 | 1-19, 20+ | 10, 26.03 | 4.47, 10.34 M |
| 54: KAISE2 | 1-19, 20+ | 10, 26.03 | 7.61, 22.12 M |
| 55: KAISER | 1-19, 20-40, 41+ | 10, 24.32, 53.33 | 6.58, 17.24, 20.91 M |
| 56: KAISER | 1-19, 20-40, 41+ | 10, 24.32, 53.33 | 3.42, 7.98, 12.63 M |
| 57: KANELL | 1-10, 11-20, 21-35, 36+ | 7.08, 17.67, 26.71, 45.71 | 1.71, 7.06, 20.39, 34.22 M |
| 58: KATSOU | 1-20, 21+ | 13.38, 31.83 | 2.26, 7.46 M |
| 59: KAUFMA | 1-14, 15-24, 25-34, 35-44, 45+ | 8.11, 19.19, 28.13, 39, 53.33 | 8.00, 15.00, 28.00, 43.00, 60.00 M |
| 60: KINLEN | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 10.61, 14.14, 21.74 M |
| 61: KNEKT | 1-14, 15+ | 8.11, 24.09 | 5.00, 12.70 M |
| 62: KOO | 1-10, 11-20, 21-30 | 7.08, 17.67, 25.88 | 1.36, 7.29, 1.52 |
| 63: LIAW | 1-10, 11-20, 21+ | 7.08, 17.67, 31.83 | 3.10, 3.60, 8.30 M |
| 64: LIDDEL | 1-19, 20+ | 10, 26.03 | 3.33, 5.02 M |
| 65: MACLEN | 1-9, 10-19, 20-29, 30+ | 4.85, 12.73, 21.35, 38.39 | 1.36, 3.41, 4.16, 5.00 M |
| 66: MACLEN | 1-9, 10-19, 20+ | 4.85, 12.73, 26.03 | 0.76, 3.44, 3.84 |
| 67: MATOS | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 1.60, 8.00, 15.00 M |

| | | | |
|---|---|---|---|
| 68: MIGRAN | 1-9, 10-19, 20, 21+ | 4.85, 12.73, 20, 31.83 | 4.01, 4.24, 5.14, 5.93 M |
| 69: MIGRAN | 1-9, 10-19, 20 | 4.85, 12.73, 20 | 4.88, 6.53, 7.48 M |
| 70: MRFITR | 1-19, 20-39, 40+ | 10, 22.90, 45.71 | 10.86, 50.12, 56.43 M |
| 71: NAM | 1-24, 25+ | 14.06, 34.63 | 6.70, 10.27 M |
| 72: NAM | 1-24, 25+ | 14.06, 34.63 | 9.06, 16.65 M |
| 73: PARKIN | 1-14, 15+ | 8.11, 24.09 | 3.90, 5.20 M |
| 74: PERSH2 | 1-9, 10+ | 4.85, 20.90 | 5.76, 11.34 M |
| 75: PETO | 1-14, 15+ | 8.11, 24.09 | 5.50, 9.49 M |
| 76: PEZZO2 | 1-20, 21-40, 41+ | 13.38, 29.02, 53.33 | 8.00, 44.39, 112.13 M |
| 77: PEZZOT | 1-20, 21-40, 41+ | 13.38, 29.02, 53.33 | 7.40, 70.00, 246.50 M |
| 78: PRESCO | 1-14, 15+ | 8.11, 24.09 | 10.20, 19.96 M |
| 79: PRESCO | 1-14, 15+ | 8.11, 24.09 | 6.36, 10.08 M |
| 80: SEGI2 | 1-9, 10-19, 20-29, 30-39, 40+ | 4.85, 12.73, 21.33, 31.07, 45.71 | 2.10, 3.10, 3.40, 6.90, 7.90 M |
| 81: SEGI2 | 1-9, 10-19, 20+ | 4.85, 12.73, 26.03 | 2.90, 1.44, 1.03 |
| 82: SHAW | 1-19, 20+ | 10, 26.03 | 6.31, 30.48 M |
| 83: SOBUE | 1-19, 20-29, 30+ | 10, 21.35, 38.39 | 3.52, 4.00, 4.55 M |
| 84: SPEIZE | 1-4, 5-14, 15-24, 25-34, 35+ | 2.5, 9.42, 19.19, 28.13, 44.38 | 2.70, 5.20, 12.60, 15.70, 22.00 M |
| 85: STOCKW | 1-19, 20-40, 41+ | 10, 24.32, 53.33 | 6.67, 14.51, 28.84 M |
| 86: SVENSS | 1-10, 11-20, 21+ | 7.08, 17.67, 31.83 | 4.60, 12.60, 59.00 M |
| 87: TENKAN | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 15.86, 20.25, 24.97 M |
| 88: TSUGAN | 1-15, 16-35, 36+ | 9.33, 22.45, 45.71 | 0.90, 1.22, 1.66 |
| 89: TULINI | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 6.02, 12.00, 27.30 M |
| 90: TULINI | 1-14, 15-24, 25+ | 8.11, 19.19, 34.63 | 8.17, 26.30, 38.70 M |
| 91: TVERDA | 1-9, 10-19, 20+ | 4.85, 12.73, 26.03 | 2.14, 3.32, 6.56 M |
| 92: TVERDA | 1-9, 20+ | 4.85, 26.03 | 4.53, 18.00 M |
| 93: WAKAI | 1-19, 20-20, 30+ | 10.00, 21.35, 38.39 | 1.80, 4.01, 9.19 M |
| 94: WU | 1-20, 21+ | 13.38, 31.83 | 3.25, 8.48 M |
| 95: WYNDE6 | 1-10, 11-20, 21-30, 31+ | 7.08, 17.67, 25.88, 43.06 | 6.80, 11.16, 17.32, 28.22 M |
| 96: WYNDE6 | 1-10, 11-20, 21-30, 31+ | 7.08, 17.67, 25.88, 38.39 | 3.75, 11.97, 21.64, 39.14 M |
| 97: YAMAGU | 1-20, 21+ | 13.38, 31.83 | 3.75, 12.14 M |

[1]In some studies, amount smoked is based on cigarette equivalents for cigars and pipes; [2]M indicates a strictly monotonic rise in RR with increasing amount smoked.

concern age-adjusted RRs. Race and/or other factors were adjusted for in 28 (28.9%) blocks.

Table 2 gives for each block the levels used to categorize amount smoked and the corresponding estimated mean values and RRs for each level. The RRs reveal an obvious trend for risk to rise with amount smoked. Of the 96 blocks with more than one level, 84 (87.5%) show a strictly monotonic increase in RR. However, considerable variation is evident in the RR for the highest exposure.

Table 3 gives the pool-first results investigating model suitability. The exponential model is particularly poor, explaining only 21.75% of the overall deviance in the estimates of log RR. The log model is also relatively poor. The linear, quadratic and cubic models are better. However, despite involving more parameters, the cubic model explains less of the overall deviance than do the best-fitting power or log-with-baseline models. The residual deviance is lowest for the log-with-baseline model, the best-fitting W value explaining 94.12% of the overall deviance, though the best-fitting power model explains almost as much (93.95%).

Fit Amount Smoked[11], gives full details for the further analyses carried out using the linear, and best-fitting power and log-with-baseline models. These include 95%CIs for the RRs in Table 2, and observed and fitted numbers by level for each block.

For each of these models, Table 4 compares the observed and fitted numbers of cases summed over blocks for never smokers and for current smokers by amount smoked. The linear model fits poorly, overestimating cases for never smokers and 30+ cigs/d smokers and underestimating for 1-30 cigs/d smokers, the model implying a far steeper increase with amount smoked than observed. This is consistent with the block-specific goodness-of-fit tests, 63 showing misfits significant at $P < 0.05$. This model is clearly inadequate for amount smoked.

Although Table 4 shows highly significant ($P < 0.001$) misfit to both the power and log-with-baseline models, the misfit is not substantial, with observed and expected numbers generally agreeing to a few percent.

For each block, and both models, Table 5 gives fitted values of $\beta_1$ and SE and goodness-of-fit $P$ values. A number of blocks show significant ($P < 0.05$) misfit, these tending to be the same blocks for both models. We comment on those 15 blocks where the $P$ value for the log-with-baseline model is $< 0.01$ (Fit Amount Smoked[11] and Table 2 for further details). These divide into various categories. Three blocks (19: CPS I , 34: DORN, 37: ENSTRO males) involve very large numbers of cases (Table 3) where the model appears to fit quite well, though in block 19: CPS I the observed flattening of response for 40+ cigs/d is not well fitted. Seven blocks (6: BEST, 20: CPS Ⅱ males, 22: DARBY males, 43: HAMMON, 60: KINLEN, 74: PERSH2, 87: TENKAN) show a marked risk increase for the lowest level of amount smoked, but the slope subsequently flattens. In contrast the reverse is true for five blocks (24: DEAN3 males, 38: ENSTRO fe-

**Table 3 Comparing the suitability of different models relating log *RR* to amount smoked by current smokers, expressed as d = cigarettes per day/10**

| Model | Parameter value[1] | Fitted coefficient(s) (SE) | Deviance | DF | Deviance explained (%) |
|---|---|---|---|---|---|
| Null | - | - | 24894.53 | 97 | |
| Linear: log RR = $\beta_1 d$ | - | $\beta_1$ = 0.6107 (0.0046) | 7265.32 | 96 | 70.82 |
| Quadratic: log RR = $\beta_1 d + \beta_2 d^2$ | - | $\beta_1$ = 1.4121 (0.0130), $\beta_2$ = -0.1792 (0.0027) | 2907.39 | 95 | 88.32 |
| Cubic: log RR = $\beta_1 d + \beta_2 d^2 + \beta_3 d^3$ | - | $\beta_1$ = 2.1915 (0.0266), $\beta_2$ = -0.6346 (0.0138), $\beta_3$ = 0.0633 (0.0019) | 1779.05 | 94 | 92.86 |
| Power: log RR = $\beta_1 d^Y$ | Y = 0.32 | $\beta_1$ = 1.8922 (0.0124) | 1512.49 | 96 | |
| | Y = 0.33 | $\beta_1$ = 1.8691 (0.0122) | 1506.19 | 96 | |
| | Y = 0.34 | $\beta_1$ = 1.8457 (0.0121) | 1506.07 | 96 | 93.95 |
| | Y = 0.35 | $\beta_1$ = 1.8222 (0.0119) | 1511.92 | 96 | |
| | Y = 0.36 | $\beta_1$ = 1.7986 (0.0118) | 1523.54 | 96 | |
| | Y = 0.50 | $\beta_1$ = 1.4673 (0.0097) | 2179.71 | 96 | |
| | Y = 1.00 | $\beta_1$ = 0.6107 (0.0046) | 7265.32 | 96 | |
| | Y = 2.00 | $\beta_1$ = 0.0969 (0.0010) | 14739.19 | 96 | |
| Log: log RR = $\beta_1 \log d$ | - | $\beta_1$ = 1.2265 (0.0107) | 11674.70 | 96 | 53.10 |
| Exponential: log RR = $\beta_1 \exp d$ | - | $\beta_1$ = 0.0120 (0.0002) | 19480.21 | 96 | 21.75 |
| Log-with-baseline: log RR = $\beta_1 \log (1 + Wd)$ | W = 7.5 | $\beta_1$ = 0.8520 (0.0056) | 1466.21 | 96 | |
| | W = 7.7 | $\beta_1$ = 0.8456 (0.0055) | 1465.37 | 96 | |
| | W = 7.9 | $\beta_1$ = 0.8394 (0.0054) | 1464.88 | 96 | |
| | W = 8.0 | $\beta_1$ = 0.8364 (0.0055) | 1464.76 | 96 | |
| | W = 8.1 | $\beta_1$ = 0.8334 (0.0054) | 1464.71 | 96 | 94.12 |
| | W = 8.2 | $\beta_1$ = 0.8305 (0.0054) | 1464.73 | 96 | |
| | W = 8.3 | $\beta_1$ = 0.8277 (0.0054) | 1464.82 | 96 | |

[1]Note that we only sought the best-fitting value of Y to two decimal places and of W to one decimal place.

**Table 4 Amount smoked by current smokers - observed and fitted lung cancers for the linear, best power and best log-with-baseline model, with $\beta_1$ fitted separately for each block**

| Midpoint amount smoked (cigs/d) | Observed[1] | Fitted[2] | | |
|---|---|---|---|---|
| | | Linear model | Best power model | Best log-with-baseline model |
| < 5 | 1249.17 | 1023.00 | 1297.42 | 1173.88 |
| 5 to < 10 | 2579.62 | 2156.31 | 2595.37 | 2539.78 |
| 10 to < 15 | 6125.69 | 4749.92 | 6276.74 | 6299.68 |
| 15 to < 20 | 7940.18 | 6678.14 | 8009.06 | 8156.50 |
| 20 to < 30 | 18138.36 | 15724.45 | 17468.99 | 17678.51 |
| 30 to < 40 | 3858.94 | 4106.12 | 3743.23 | 3701.31 |
| 40+ | 7703.88 | 8860.42 | 8115.77 | 7949.95 |
| Never smoked | 6649.61 | 10947.08 | 6738.86 | 6745.84 |
| Total | 54245.45 | 54245.45 | 54245.45 | 54245.45 |
| Fit statistic[3] | | 3792.53 | 65.29 | 56.78 |

[1]Observed pseudo-number of lung cancer cases, summed over blocks; [2]Fitted pseudo-number of lung cancer cases, summed over blocks. For each study, the fitted number of cases for each block is calculated from the numbers at risk and the fitted RRs by amount smoked, derived from the fitted value of $\beta_1$; [3]Based on summation of (observed-fitted)$^2$/fitted, the summation also including terms for the observed and fitted total numbers of controls (not shown). The statistic can be considered to be approximately chisquared on 12 DF and is significant at $P < 0.001$ for all three models.

males, 57: KANELL, 76: PEZZO2, 77: PEZZOT) with the RR for the highest exposure greater than predicted from the response at lower levels. For some of the 15 blocks, the number of cases in never smokers is relatively low (less than 10 in 6 of them) and the best-fitting model gives rather different fitted numbers, so the fitted block of RRs appears substantially different from that observed For example, in block 6: BEST where the observed pseudo-number of cases in never smokers is 6.88, and the observed RRs are 10.00, 16.41 and 17.31 for 1-9, 10-20 and 21+ cigs/d, the fitted number of cases in never smokers is 23.61 and the fitted RRs are 2.47, 4.46 and 6.47.

Table 6 presents results of weighted simple regression analyses of $\beta_1$ for the log-with-baseline model. There is highly significant ($P < 0.001$) variation by continent, with $\beta_1$ much lower for Asian studies, and by study size, larger studies giving higher $\beta_1$ values. Some variation is also seen for sex ($P < 0.05$), study type, publication year and midpoint age ($P < 0.1$), but not with product definition or unexposed group. Table 6 also presents predicted RRs at 20 cigs/d. The variation by continent is clear.

In a forward stepwise analysis (not shown), continent remained highly significant ($P < 0.001$), but no other factor remained significant at $P < 0.05$. The association with study size seems due to a strong correlation with continent.

**Table 5 Amount smoked by current smokers - fitted values of $\beta_1$ and SE, and *P* values for goodness-of-fit tests for the best-fitting power model and log-with-baseline model**

| Block: Study | Log-with-baseline model[1] log RR = $\beta_1$ [1 + (8.10c/10)] | | | Power model[1] log RR = $\beta_1$ [(c/10)$^{0.34}$] | | |
|---|---|---|---|---|---|---|
| | $\beta_1$ | SE $\beta_1$ | *P* (fit)[2] | $\beta_1$ | SE $\beta_1$ | *P* (fit)[2] |
| 1: AKIBA | 0.6599 | 0.0712 | NS | 1.4882 | 0.1609 | NS |
| 2: AKIBA | 0.6099 | 0.0660 | NS | 1.3445 | 0.1453 | NS |
| 3: ARCHER | 0.6713 | 0.1274 | NS | 1.4932 | 0.2835 | NS |
| 4: AXELSS | 1.3245 | 0.2214 | NS | 2.9770 | 0.4976 | NS |
| 5: BENSHL | 0.7133 | 0.1169 | NS | 1.6337 | 0.2688 | NS |
| 6: BEST | 0.5678 | 0.0824 | 0.0000 | 1.3483 | 0.1948 | 0.0001 |
| 7: BOUCOT | 0.8332 | 0.2140 | NS | 1.7439 | 0.4482 | NS |
| 8: BRETT | 0.6637 | 0.1146 | NS | 1.4946 | 0.2575 | NS |
| 9: BROSS | 0.6090 | 0.0735 | NS | 1.3514 | 0.1634 | NS |
| 10: BUFFLE | 0.9349 | 0.1040 | NS | 2.0810 | 0.2313 | NS |
| 11: CEDERL | 0.8191 | 0.0726 | NS | 1.8388 | 0.1634 | NS |
| 12: CEDERL | 0.7697 | 0.0693 | NS | 1.6833 | 0.1520 | NS |
| 13: CHANG | 0.6772 | 0.1522 | NS | 1.5044 | 0.3402 | NS |
| 14: CHANG | 0.6255 | 0.1165 | NS | 1.3937 | 0.2595 | NS |
| 15: CHOW | 1.0002 | 0.1050 | NS | 2.2134 | 0.2311 | NS |
| 16: COMSTO | 0.8497 | 0.1552 | NS | 1.8491 | 0.3399 | NS |
| 17: COMSTO | 0.8857 | 0.1172 | NS | 1.9560 | 0.2601 | NS |
| 18: CORREA | 0.9945 | 0.0483 | NS | 2.2052 | 0.1069 | NS |
| 19: CPS I | 0.8314 | 0.0348 | 0.0002 | 1.8278 | 0.0773 | 0.0000 |
| 20: CPS II | 0.8262 | 0.0331 | 0.0000 | 1.8467 | 0.0739 | 0.0000 |
| 21: CPS II | 0.8972 | 0.0304 | 0.0153 | 1.9874 | 0.0677 | 0.0006 |
| 22: DARBY | 0.8198 | 0.1213 | 0.0000 | 1.9085 | 0.2770 | 0.0000 |
| 23: DARBY | 1.0879 | 0.0841 | NS | 2.4434 | 0.1879 | NS |
| 24: DEAN3 | 0.8271 | 0.0724 | 0.0009 | 1.8968 | 0.1639 | 0.0042 |
| 25: DEAN3 | 0.8617 | 0.0801 | (0.0569) | 1.9152 | 0.1789 | 0.0368 |
| 26: DEKLER | 0.5704 | 0.1839 | NS | 1.3114 | 0.4155 | NS |
| 27: DOLL2 | 1.0464 | 0.0644 | (0.0824) | 2.3531 | 0.1443 | NS |
| 28: DOLL2 | 1.1001 | 0.1698 | NS | 2.4424 | 0.3745 | 0.0091 |
| 29: DORANT | 1.1082 | 0.0867 | 0.0322 | 2.5183 | 0.1972 | 0.0270 |
| 30: DORGAN | 0.9728 | 0.0937 | NS | 2.1814 | 0.2101 | NS |
| 31: DORGAN | 1.4168 | 0.1835 | NS | 3.1557 | 0.4086 | NS |
| 32: DORGAN | 0.9636 | 0.0595 | NS | 2.1418 | 0.1325 | NS |
| 33: DORGAN | 1.0359 | 0.1844 | NS | 2.2976 | 0.4094 | NS |
| 34: DORN | 0.9024 | 0.0171 | 0.0001 | 2.0099 | 0.0381 | 0.0000 |
| 35: ENGELA | 1.0083 | 0.1270 | NS | 2.3043 | 0.2995 | NS |
| 36: ENGELA | 1.0897 | 0.1561 | 0.0206 | 2.5654 | 0.3553 | NS |
| 37: ENSTRO | 0.8261 | 0.0312 | 0.0000 | 1.7910 | 0.0684 | 0.0000 |
| 38: ENSTRO | 0.8229 | 0.0252 | 0.0000 | 1.8368 | 0.0564 | 0.0000 |
| 39: GAO2 | 0.7037 | 0.1042 | NS | 1.5526 | 0.2304 | NS |
| 40: GILLIS | 0.5811 | 0.0729 | NS | 1.2704 | 0.1614 | NS |
| 41: HAENSZ | 0.3382 | 0.0805 | NS | 0.7563 | 0.1796 | NS |
| 42: HAMMO2 | 0.5198 | 0.1167 | 0.0112 | 1.1870 | 0.2641 | 0.0138 |
| 43: HAMMON | 0.8032 | 0.0748 | 0.0038 | 1.8185 | 0.1677 | 0.0116 |
| 44: HIRAYA | 0.5974 | 0.0337 | NS | 1.3374 | 0.0756 | NS |
| 45: HIRAYA | 0.4424 | 0.0501 | NS | 0.9729 | 0.1100 | NS |
| 46: HITOSU | 0.4573 | 0.1182 | NS | 1.0324 | 0.2654 | NS |
| 47: HITOSU | 0.4988 | 0.1223 | NS | 1.0992 | 0.2686 | NS |
| 48: HOLE | 0.6142 | 0.1082 | NS | 1.3410 | 0.2413 | NS |
| 49: HUMBLE | 1.0435 | 0.1431 | NS | 2.3386 | 0.3207 | NS |
| 50: HUMBLE | 1.0378 | 0.2633 | NS | 2.3306 | 0.5907 | NS |
| 51: HUMBLE | 0.8922 | 0.1393 | NS | 1.9997 | 0.3114 | NS |
| 52: HUMBLE | 1.2268 | 0.2487 | NS | 2.7318 | 0.5532 | NS |
| 53: KAISE2 | 0.7599 | 0.1018 | NS | 1.6991 | 0.2278 | NS |
| 54: KAISE2 | 1.0098 | 0.1107 | NS | 2.2592 | 0.2478 | NS |
| 55: KAISER | 0.7685 | 0.0477 | 0.0247 | 1.6130 | 0.1010 | 0.0033 |
| 56: KAISER | 0.6882 | 0.0570 | NS | 1.4932 | 0.1238 | NS |
| 57: KANELL | 0.8982 | 0.0647 | 0.0000 | 1.9758 | 0.1442 | 0 |
| 58: KATSOU | 0.4517 | 0.1260 | NS | 1.0089 | 0.2809 | NS |
| 59: KAUFMA | 1.0712 | 0.0583 | NS | 2.3560 | 0.1278 | NS |
| 60: KINLEN | 0.6021 | 0.0616 | 0.0017 | 1.3838 | 0.1399 | 0.0055 |
| 61: KNEKT | 0.8644 | 0.1264 | NS | 1.9618 | 0.2876 | NS |
| 62: KOO | 0.4309 | 0.1424 | NS | 0.9299 | 0.3133 | NS |
| 63: LIAW | 0.5538 | 0.0918 | NS | 1.2384 | 0.2045 | NS |
| 64: LIDDEL | 0.5136 | 0.0742 | NS | 1.1535 | 0.1666 | NS |

| | | | | | | |
|---|---|---|---|---|---|---|
| 65: MACLEN | 0.4973 | 0.1511 | NS | 1.0875 | 0.3355 | NS |
| 66: MACLEN | 0.4301 | 0.1161 | NS | 0.9350 | 0.2577 | NS |
| 67: MATOS | 0.8315 | 0.1122 | NS | 1.8390 | 0.2485 | NS |
| 68: MIGRAN | 0.4590 | 0.1448 | NS | 1.0568 | 0.3286 | NS |
| 69: MIGRAN | 0.7246 | 0.2207 | NS | 1.6547 | 0.4987 | NS |
| 70: MRFITR | 0.6382 | 0.2225 | 0.0338 | 1.1984 | 0.4461 | 0.0152 |
| 71: NAM | 0.6883 | 0.0711 | NS | 1.5157 | 0.1569 | NS |
| 72: NAM | 0.8481 | 0.0690 | NS | 1.8809 | 0.1532 | NS |
| 73: PARKIN | 0.6129 | 0.0554 | NS | 1.3556 | 0.1222 | NS |
| 74: PERSH2 | 0.8420 | 0.0381 | 0.0004 | 1.9065 | 0.0854 | 0.0446 |
| 75: PETO | 0.6262 | 0.1606 | NS | 1.4651 | 0.3737 | NS |
| 76: PEZZO2 | 1.3641 | 0.1251 | 0.0051 | 2.9784 | 0.2710 | 0.0118 |
| 77: PEZZOT | 1.5483 | 0.1569 | 0.0005 | 3.4045 | 0.3415 | 0.0014 |
| 78: PRESCO | 0.8289 | 0.1002 | (0.0759) | 1.9305 | 0.2318 | NS |
| 79: PRESCO | 0.7383 | 0.0942 | NS | 1.6725 | 0.2122 | NS |
| 80: SEGI2 | 0.5679 | 0.0991 | NS | 1.2849 | 0.2208 | NS |
| 81: SEGI2 | 0.1503 | 0.1297 | NS | 0.3513 | 0.2865 | NS |
| 82: SHAW | 1.1326 | 0.1121 | NS | 2.5298 | 0.2509 | NS |
| 83: SOBUE | 0.4048 | 0.0594 | NS | 0.8893 | 0.1309 | NS |
| 84: SPEIZE | 0.8772 | 0.0390 | NS | 1.9509 | 0.0870 | NS |
| 85: STOCKW | 0.8822 | 0.0096 | NS | 1.9383 | 0.0210 | NS |
| 86: SVENSS | 0.9255 | 0.1158 | NS | 2.0479 | 0.2573 | NS |
| 87: TENKAN | 0.6211 | 0.1092 | 0.0001 | 1.4407 | 0.2484 | 0.0003 |
| 88: TSUGAN | 0.1104 | 0.1169 | NS | 0.2461 | 0.2574 | NS |
| 89: TULINI | 1.0019 | 0.0901 | NS | 2.2394 | 0.2008 | NS |
| 90: TULINI | 1.2489 | 0.0862 | (0.0728) | 2.8447 | 0.1983 | 0.0098 |
| 91: TVERDA | 0.6063 | 0.0773 | NS | 1.3657 | 0.1746 | NS |
| 92: TVERDA | 0.9309 | 0.1838 | NS | 2.1244 | 0.4198 | NS |
| 93: WAKAI | 0.6776 | 0.1099 | (0.0831) | 1.5073 | 0.2430 | NS |
| 94: WU | 0.5916 | 0.1017 | NS | 1.3183 | 0.2263 | NS |
| 95: WYNDE6 | 0.9181 | 0.0373 | NS | 2.0237 | 0.0820 | NS |
| 96: WYNDE6 | 0.9796 | 0.0371 | 0.0102 | 2.1767 | 0.0825 | 0.0060 |
| 97: YAMAGU | 0.6608 | 0.1223 | NS | 1.4756 | 0.2723 | NS |

[1]c = cigarettes/d; [2]Not significant (NS) indicates $P \geq 0.1$. $P$ values in the range $0.05 \leq P < 0.1$ are shown in brackets.

**Table 6 Amount smoked by current smokers - inverse-variance weighted simple regression analyses of $\beta_1$ based on best-fitting log + baseline model**

| Factor | Level | *n* | $\beta_1$ (95% CI) | *P*[1] | RR for 20 cigs/d |
|---|---|---|---|---|---|
| All | | 97 | 0.83 (0.80-0.86) | | 10.71 |
| Sex | Male | 55 | 0.79 (0.75-0.84) | < 0.05 | 9.50 |
| | Female | 34 | 0.82 (0.75-0.88) | | 10.21 |
| | Combined | 8 | 0.89 (0.84-0.94) | | 12.50 |
| Study type | Case-control | 49 | 0.86 (0.82-0.90) | < 0.1 | 11.48 |
| | Prospective | 48 | 0.80 (0.76-0.85) | | 9.85 |
| Continent | North America | 43 | 0.87 (0.84-0.89) | < 0.001 | 11.48 |
| | Europe | 32 | 0.83 (0.76-0.90) | | 10.65 |
| | Asia | 17 | 0.53 (0.45-0.61) | | 4.53 |
| | Other | 5 | 0.80 (0.61-0.99) | | 9.80 |
| Publication year[2] | < 1990 | 40 | 0.83 (0.79-0.88) | < 0.1 | 10.71 |
| | 1990-1994 | 29 | 0.77 (0.70-0.84) | | 8.90 |
| | 1995-1999 | 28 | 0.87 (0.82-0.92) | | 11.94 |
| Product[3] | Any product | 20 | 0.75 (0.64-0.85) | NS | 8.36 |
| | Cigarettes +/- | 61 | 0.85 (0.81-0.88) | | 11.09 |
| | Cigarettes only | 16 | 0.82 (0.74-0.90) | | 10.19 |
| Unexposed | Never cigarettes | 32 | 0.83 (0.76-0.90) | NS | 10.49 |
| | Never any product | 65 | 0.84 (0.80-0.87) | | 10.76 |
| Grouped midpoint age (yr) | < 50 | 14 | 0.79 (0.67-0.92) | < 0.1 | 9.57 |
| | 50-59 | 62 | 0.85 (0.82-0.89) | | 11.35 |
| | 60+ | 21 | 0.79 (0.71-0.84) | | 9.13 |
| Cases in smokers | < 100 | 29 | 0.67 (0.56-0.78) | < 0.001 | 6.72 |
| | 100 to < 200 | 28 | 0.73 (0.64-0.83) | | 8.08 |
| | 200 to < 500 | 16 | 0.76 (0.65-0.87) | | 8.72 |
| | 500 to < 1000 | 16 | 0.80 (0.75-0.86) | | 9.87 |
| | 1000+ | 8 | 0.89 (0.85-0.92) | | 12.42 |

[1]Probability values for factor considered, presented as < 0.001, < 0.01, < 0.05, < 0.1 or not significant (NS) ($P \geq 0.1$); [2]Of principal publication for the study; [3]Any product = Smokes cigarettes and/or pipes and/or cigars; Cigarettes +/- = Smokes cigarettes with or without other products (pipes, cigars).

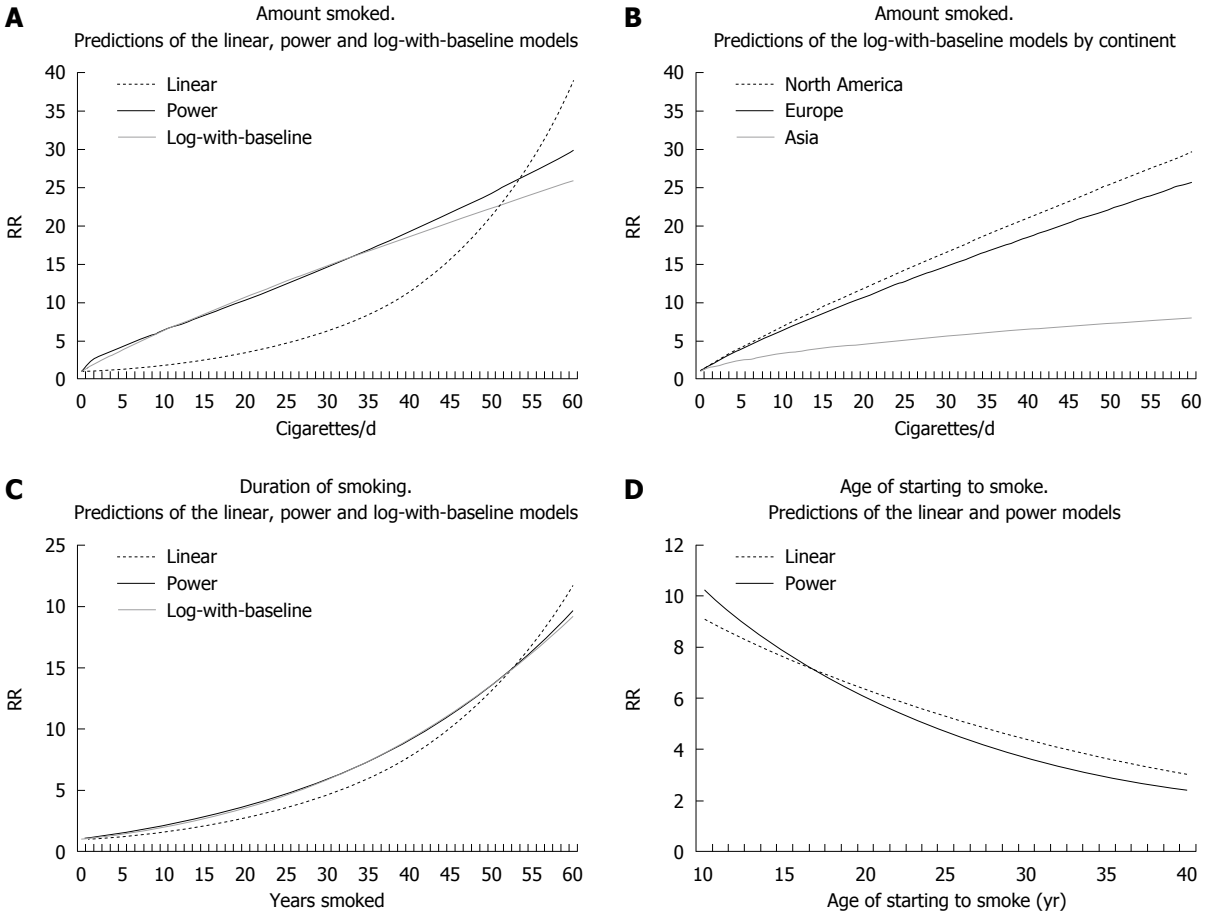



**Figure 1 Model predictions.** A: Amount smoked. Predictions of the linear, power and log-with-baseline models; B: Amount smoked. Predictions of the log-with-baseline models by continent; C: Duration of smoking. Predictions of the linear, power, and log-with-baseline models; D: Age of starting to smoke. Predictions of the linear and power models. The RR is plotted against number of cigarettes per day (A and B), years smoked (C) and age of starting to smoke (D). The linear model for amount smoked is the poorest fit to the data, other models shown fitting the data similarly well.

For the 97 blocks combined the RRs predicted by the log-with-baseline model for 5, 10, 20, 30, 40 and 50 cigs/d are, respectively, 3.86, 6.30, 10.71, 14.77, 18.62 and 22.31. The RRs are similar for the power model (4.30, 6.33, 10.34, 14.61, 19.24 and 24.29), but very different for the linear model (1.36, 1.84, 3.39, 6.25, 11.50 and 21.19). See also Figure 1A which compares model predictions, and Figure 1B which shows the predictions by continent.

### *Duration of smoking*

Blocks[11] gives details of the 35 blocks, derived from 14 studies. CPS I and CPS Ⅱ provide 20 blocks, with data by sex and age. Three further studies provide sex-specific results, the remaining nine only providing one block each. Three studies are from Europe, two South America and one Asia, the remaining eight being from North America.

Table 7 summarizes the dose-response data. A clear increase in risk with increasing duration is evident, 34 blocks (97.1%) showing a greater RR for the longest than the shortest duration group, and 22 (62.9%) showing a strictly monotonic increase in RR.

Table 8 summarizes the analyses on model suitability. The exponential model is again very poor, explaining only 31.47% of the deviance. Other models differ little, explaining 88.49% to 89.96%. The best single parameter models are the best-fitting power model (89.96%) and log-with-baseline model (89.89%).

Fit Duration[11], gives full details for the further analyses using the linear model and best-fitting power and log-with-baseline models, laid out as Fit Amount Smoked[11].

Table 9 compares observed and fitted cases summed over blocks. Misfit is similar for all three models, and significant ($P < 0.001$), though its extent seems relatively moderate.

Table 10 gives fitted values of $\beta_1$ and SE and goodness-of-fit $P$ values for the power and log-with-baseline models. We comment on six blocks where $P$ is $< 0.001$ for both models, and three where $P$ is $< 0.001$ for one model (Fit Duration[11] and Table 7 for further details). In three blocks (1: AMANDU, 29: HUMBLE, 35: PEZZO2) the RR associated with the lowest duration level is large, but the RRs associated with higher levels are not much larger (or even smaller). In another block (10: CPS I males age 65-74 years) the misfit comes from the lack of rise in risk over short duration levels, while in another (17: CPS Ⅱ males age 30-44 years) it is associated with the very high risk for the longest duration. In three other blocks (11: CPS I males age 75+ years, 15: CPS I females age 65-74 years, 20: CPS Ⅱ males age 65-74 years) the misfit is at least partly due to the non-monotonic dose-response.

**Table 7 Duration of smoking by current smokers (yr) - dose-response data**

| Block: Study | Duration of smoking groupings (yr) | Mean values | RRs[1] |
|---|---|---|---|
| 1: AMANDU | 0-24, 25+ | 12.66, 41.28 | 5.92, 7.02 M |
| 2: BEST | 1-4, 5-9, 10-14, 15-19, 20-29, 30-39, 40+ | 2.64, 7.02, 12.03, 16.89, 24.03, 33.93, 51.12 | 1.60, 2.60, 2.30, 3.20, 4.10, 13.90, 14.20 |
| 3: BOUCOT | 1-39, 40+ | 32.05, 51.14 | 42.40, 89.94 M |
| 4: BUFFLE | 1-30, 31-40, 41+ | 19.34, 35.56, 48.49 | 14.00, 14.70, 18.15 M |
| 5: CEDERL | 1-29, 30+ | 23.94, 41.09 | 1.80, 7.40 M |
| 6: CEDERL | 1-29, 30+ | 22.26, 39.47 | 1.60, 9.60 M |
| 7: CPS I | 1-29, 30+ | 23.70, 32.03 | 3.83, 6.56 M |
| 8: CPS I | 1-29, 30-34, 35-39, 40+ | 24.92, 31.91, 36.52, 43.07 | 5.58, 13.40, 16.93, 29.61 M |
| 9: CPS I | 1-29, 30-34, 35-39, 40-44, 45-49, 50+ | 23.57, 32.25, 37.75, 42.03, 46.80, 51.81 | 3.11, 6.57, 10.35, 15.21, 21.46, 33.11 M |
| 10: CPS I | 1-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60+ | 10.67, 30.00, 39.00, 42.21, 47.57, 52.05, 56.73, 62.87 | 5.46, 6.26, 5.86, 8.22, 12.48, 15.28, 18.86, 28.60 |
| 11: CPS I | 1-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60+ | 20.00, 32.00, 37.00, 41.00, 46.25, 51.86, 57.00, 65.03 | 2.17, 2.27, 12.66, 3.47, 6.22, 6.31, 12.87, 13.23 |
| 12: CPS I | 1-29, 30+ | 21.86, 31.37 | 6.16, 13.80 M |
| 13: CPS I | 1-29, 30-34, 35-39, 40+ | 23.29, 31.61, 36.72, 43.09 | 2.90, 6.91, 8.49, 19.48 M |
| 14: CPS I | 1-29, 30-34, 35-39, 40-44, 45-49, 50+ | 21.03, 31.92, 37.45, 42.21, 46.65, 52.80 | 1.51, 3.71, 4.73, 5.72, 7.78, 14.48 M |
| 15: CPS I | 1-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55+ | 21.69, 31.38, 37.19, 41.80, 47.14, 51.76, 57.56 | 2.30, 3.38, 3.67, 5.81, 7.01, 6.20, 3.32 |
| 16: CPS I | 1-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55+ | 18.00, 30.00, 38.00, 41.75, 46.71, 52.25, 61.54 | 0.38, 3.11, 2.72, 3.22, 1.15, 1.85, 3.20 |
| 17: CPS II | 1-29, 30+ | 19.15, 32.13 | 3.69, 108.27 M |
| 18: CPS II | 1-29, 30-34, 35-39, 40+ | 24.92, 31.91, 36.52, 43.07 | 7.72, 19.47, 25.01, 36.98 M |
| 19: CPS II | 1-29, 30-34, 35-39, 40-44, 45-49, 50+ | 23.57, 32.25, 37.75, 42.03, 46.80, 51.81 | 15.51, 17.86, 27.31, 44.71, 50.92, 72.07 M |
| 20: CPS II | 1-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60+ | 10.67, 30.00, 39.00, 42.21, 47.57, 52.05, 56.73, 62.87 | 9.97, 11.05, 20.45, 18.59, 24.33, 32.06, 40.43, 45.64 |
| 21: CPS II | 1-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60+ | 20.00, 32.00, 37.00, 41.00, 46.25, 51.86, 57.00, 66.09 | 7.54, 4.53, 4.37, 16.30, 16.01, 13.40, 18.09, 21.79 |
| 22: CPS II | 1-19, 20+ | 13.93, 24.23 | 10.98, 8.38 |
| 23: CPS II | 1-29, 30-34, 35+ | 23.29, 31.61, 37.81 | 8.96, 18.34, 23.32 M |
| 24: CPS II | 1-29, 30-34, 35-39, 40-44, 45-49, 50+ | 21.03, 31.92, 37.45, 42.22, 46.65, 52.80 | 6.50, 13.04, 17.07, 20.56, 23.94, 28.45 M |
| 25: CPS II | 1-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55+ | 21.69, 31.38, 37.19, 41.80, 47.14, 51.76, 57.56 | 6.64, 5.57, 10.19, 12.96, 15.79, 18.75, 19.34 |
| 26: CPS II | 1-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60+ | 18.00, 30.00, 38.00, 42.20, 46.71, 52.23, 57.10, 65.06 | 3.36, 7.92, 6.31, 8.14, 11.68, 9.45, 9.18, 19.88 |
| 27: DEAN2 | 1-19, 20+ | 15.09, 38.28 | 3.21, 3.84 M |
| 28: DEAN2 | 1-19, 20+ | 13.93, 35.85 | 0.98, 5.88 |
| 29: HUMBLE | 1-29, 30-39, 40-49, 50+ | 17.18, 34.07, 44.43, 56.60 | 15.45, 17.54, 19.61, 17.27 |
| 30: KAISE2 | 1-39, 40+ | 23.76, 51.12 | 4.86, 15.64 M |
| 31: KAISE2 | 1-39, 40+ | 22.73, 48.96 | 9.09, 30.41 M |
| 32: KATSOU | 1-29, 30+ | 14.58, 42.43 | 1.29, 7.43 M |
| 33: LIAW | 1-20, 21-30, 31+ | 14.32, 26.18, 44.79 | 0.90, 2.60, 4.70 |
| 34: MATOS | 1-24, 25-39, 40+ | 12.75, 31.16, 50.57 | 5.20, 7.40, 10.20 M |
| 35: PEZZO2 | 1-35, 36+ | 17.15, 48.89 | 16.25, 26.77 M |

[1]M indicates a strictly monotonic rise in RR with increasing duration of smoking.

In the remaining block (9: CPS I males age 55-64 years) the rise is monotonic, but the relatively small RR for the shortest duration does not fit in well with the large RRs for longer durations.

Table 11 presents results of weighted simple regression analysis of $\beta_1$ for the power model. Significance at $P < 0.05$ is only seen for midpoint age, with higher $\beta_1$ values for lower ages. In forward stepwise regressions (not shown), the model included in succession midpoint age ($P < 0.05$), number of cases in smokers ($P < 0.05$), smoking product ($P < 0.05$) and unexposed group ($P < 0.01$). The final model associated increased risk with younger age, larger numbers of cases, smoking of cigarettes only or cigarettes ± other products (compared to smoking any product) and with the unexposed being never any product (rather than never cigarettes).

For the 35 blocks combined the RRs predicted by the power model for durations of 10, 20, 30, 40 and 50 years are, respectively, 2.21, 3.75, 5.96, 9.11 and 13.54. Figure 1C compares model predictions.

### Age of starting to smoke

Blocks[11] gives details of the 27 blocks, deriving from 15 studies. One study gives results by age and sex, two by age for males, and five by sex but not age, the remaining seven studies only providing one block each. There are similar numbers of blocks from North America (9), Europe (9) and Asia (8), only one being from elsewhere.

Table 12 summarizes the dose-response data. A relationship of risk to age of starting is not consistently seen. While 13 blocks (48.1%) show a strictly monotonic decline in risk as starting age increases, with risk often substantially higher in early starters, and 6 (22.2%) blocks show a similar non-monotonic tendency, 8 (29.6%) blocks (1, 2, 7-9, 16, 20, 21) show no such tendency.

Table 13 summarizes the analyses on model suitability,

**Table 8 Comparing the suitability of different models relating log RR to duration of smoking by current smokers, expressed as d = years smoked/10**

| Model | Parameter value[1] | Fitted coefficient(s) (SE) | Deviance | DF | Deviance explained (%) |
|---|---|---|---|---|---|
| Null | - | - | 6161.11 | 35 | |
| Linear: log RR = $\beta_1 d$ | - | $\beta_1$ = 0.5134 (0.0069) | 672.01 | 34 | 89.09 |
| Quadratic: log RR = $\beta_1 d + \beta_2 d^2$ | - | $\beta_1$ = 0.6718 (0.0237), $\beta_2$ = -0.0300 (0.0043) | 623.20 | 33 | 89.88 |
| Cubic: log RR = $\beta_1 d + \beta_2 d^2 + \beta_3 d^3$ | - | $\beta_1$ = 0.6788 (0.0607), $\beta_2$ = -0.0330 (0.0243), $\beta_3$ = 0.0003 (0.0025) | 623.19 | 32 | 89.89 |
| Power: log RR = $\beta_1 d^Y$ | Y = 0.50 | $\beta_1$ = 1.1576 (0.0156) | 682.56 | 34 | |
| | Y = 0.72 | $\beta_1$ = 0.8180 (0.0110) | 619.30 | 34 | |
| | Y = 0.73 | $\beta_1$ = 0.8049 (0.0108) | 618.93 | 34 | |
| | Y = 0.74 | $\beta_1$ = 0.7919 (0.0106) | 618.75 | 34 | 89.96 |
| | Y = 0.75 | $\beta_1$ = 0.7791 (0.0105) | 618.76 | 34 | |
| | Y = 0.76 | $\beta_1$ = 0.7665 (0.0103) | 618.95 | 34 | |
| | Y = 1.00 | $\beta_1$ = 0.5134 (0.0069) | 672.01 | 34 | |
| | Y = 2.00 | $\beta_1$ = 0.0863 (0.0013) | 1426.44 | 34 | |
| Log: log RR = $\beta_1 \log d$ | - | $\beta_1$ = 1.5870 (0.0215) | 708.96 | 35 | 88.49 |
| Exponential: log RR = $\beta_1 \exp d$ | | $\beta_1$ = 0.0045 (0.0001) | 4222.01 | 35 | 31.47 |
| Log-with-baseline: log RR = $\beta_1 \log (1 + Wd)$ | W = 0.10 | $\beta_1$ = 6.4054 (0.0861) | 629.58 | 34 | |
| | W = 0.15 | $\beta_1$ = 4.6503 (0.0625) | 623.99 | 34 | |
| | W = 0.18 | $\beta_1$ = 4.0572 (0.0545) | 622.82 | 34 | |
| | W = 0.19 | $\beta_1$ = 3.9001 (0.0524) | 622.69 | 34 | |
| | W = 0.20 | $\beta_1$ = 3.7581 (0.0505) | 622.67 | 34 | 89.89 |
| | W = 0.21 | $\beta_1$ = 3.6293 (0.0488) | 622.74 | 34 | |
| | W = 0.22 | $\beta_1$ = 3.5117 (0.0472) | 622.89 | 34 | |

[1]Note that we only sought the best-fitting value of W and Y to two decimal places.

**Table 9 Duration of smoking by current smokers - observed and fitted lung cancers for the linear, best power and best log-with-baseline model, with $\beta_1$ fitted separately for each block**

| Midpoint years smoked | Observed[1] | Fitted[2] | | |
|---|---|---|---|---|
| | | Linear model | Best power model | Best log-with-baseline model |
| < 15 | 109.36 | 89.10 | 93.61 | 91.19 |
| 15 to < 30 | 605.64 | 584.79 | 666.12 | 645.42 |
| 30 to < 45 | 3968.47 | 3917.11 | 4000.71 | 4012.22 |
| 45+ | 3493.56 | 3573.95 | 3499.85 | 3499.95 |
| Never smoked | 2533.97 | 2546.03 | 2450.69 | 2462.20 |
| Total | 10710.98 | 10710.98 | 10710.98 | 10710.98 |
| Fit statistic[3] | | 28.14 | 25.70 | 24.53 |

[1]Observed pseudo-number of lung cancer cases, summed over blocks; [2]Fitted pseudo-number of lung cancer cases, summed over blocks. For each study, the fitted number of cases for each block is calculated from the numbers at risk and the fitted RRs by years smoked, derived from the fitted value of $\beta_1$; [3]Based on summation of (observed-fitted)$^2$/fitted, the summation also including terms for the observed and fitted total numbers of controls (not shown). The statistic can be considered to be approximately chisquared distributed on 6 df, and is significant at $P < 0.001$ for all three models.

taking (70-age at starting) as a duration-like measure. The exponential model again is the poorest, explaining only 60.24% of the deviance, and the log model is also poorer than the other models. Although slightly more deviance is explained by the cubic model, the best-fitting power model is the best simple model, explaining 88.29% of the deviance. Log-with-baseline models were also tried (not shown), but the best-fitting value of W was extremely low, so it became essentially identical to the linear model, having the same deviance.

Fit Age Start[11], gives full details of the further analyses using the linear and best-fitting power models.

Table 14 compares observed and fitted cases summed over blocks. Both the linear and best-fitted power models fit well.

Table 15 gives fitted values of $\beta_1$ and SE and goodness-of-fit *P* values. For no block does either model show misfit at $P < 0.01$, though misfits at $P < 0.05$ are sometimes seen. We comment on four blocks with some evidence ($P < 0.1$) of misfit to both models (Table 12 and Fit Age Start[11] for further details). For block 4 (CPS I males aged 55-69) the misfit seems due to the relatively small decline in risk between age of start 1-14 and 15-19 years, compared to a greater decline subsequently. For block 5 (CPS I males aged 70-84 years), risk again declines substantially over ages of start 15-19, 20-24 and 25+ years, but risk is slightly less at age 1-14 than 15-19 years. For block 12 (ENGELA males), risk decreases from age 1-19 to 20-29 years but then falls no further. For block 17 (LIAW), the pattern is non-monotonic.

Table 16 presents results of the weighted simple regression analyses of $\beta_1$ for the power model. Various fac-

**Table 10 Duration of smoking by current smokers - fitted values for of β1 and SE, and *P* values for goodness-of-fit tests for the best-fitting power model and log-with-baseline model**

| Block: Study | Log-with-baseline model[1] log RR = β1 [1 + (0.2 yr/10)] | | | Power model[1] log RR = β1 (yr/10)$^{0.74}$ | | |
|---|---|---|---|---|---|---|
| | β1 | SE β1 | *P* (fit)[2] | β1 | SE β1 | *P* (fit)[2] |
| 1: AMANDU | 1.5114 | 0.5504 | 0.0030 | 0.3454 | 0.1210 | 0.0046 |
| 2: BEST | 3.7163 | 0.4210 | (0.0548) | 0.7995 | 0.0908 | 0.0469 |
| 3: BOUCOT | 6.7731 | 1.6574 | NS | 1.4235 | 0.3482 | NS |
| 4: BUFFLE | 4.5141 | 0.4888 | 0.0466 | 0.9570 | 0.1029 | (0.0780) |
| 5: CEDERL | 3.4044 | 0.7166 | NS | 0.7147 | 0.1512 | NS |
| 6: CEDERL | 3.3566 | 0.7692 | NS | 0.6990 | 0.1609 | NS |
| 7: CPS I | 3.6386 | 1.4531 | NS | 0.7530 | 0.3009 | NS |
| 8: CPS I | 5.7248 | 0.4753 | (0.0947) | 1.2122 | 0.1011 | (0.0518) |
| 9: CPS I | 4.8058 | 0.2032 | 0.0003 | 1.0130 | 0.0429 | 0.0002 |
| 10: CPS I | 3.7681 | 0.1557 | 0.0002 | 0.7976 | 0.0328 | 0.0010 |
| 11: CPS I | 3.1437 | 0.1768 | 0.0079 | 0.6566 | 0.0369 | 0.0073 |
| 12: CPS I | 5.1328 | 1.5211 | NS | 1.0518 | 0.3119 | NS |
| 13: CPS I | 3.9960 | 0.3377 | 0.0308 | 0.8342 | 0.0708 | 0.0165 |
| 14: CPS I | 2.8575 | 0.1844 | (0.0637) | 0.5999 | 0.0388 | 0.0383 |
| 15: CPS I | 2.6384 | 0.1872 | 0.0000 | 0.5544 | 0.0394 | 0.0000 |
| 16: CPS I | 1.5443 | 0.3346 | NS | 0.3238 | 0.0702 | NS |
| 17: CPS II | 9.5482 | 2.4662 | 0.0149 | 1.9485 | 0.5160 | 0.0076 |
| 18: CPS II | 6.1160 | 0.5675 | NS | 1.2874 | 0.1199 | NS |
| 19: CPS II | 5.9905 | 0.3222 | (0.0749) | 1.2722 | 0.0684 | (0.0921) |
| 20: CPS II | 4.5913 | 0.1996 | 0.0058 | 0.9714 | 0.0421 | 0.0210 |
| 21: CPS II | 3.6604 | 0.2069 | NS | 0.7645 | 0.0432 | NS |
| 22: CPS II | 5.1073 | 3.0925 | NS | 1.1011 | 0.6362 | NS |
| 23: CPS II | 5.8103 | 0.4709 | NS | 1.2180 | 0.0988 | NS |
| 24: CPS II | 4.9498 | 0.2256 | NS | 1.0461 | 0.0476 | NS |
| 25: CPS II | 4.1028 | 0.1646 | NS | 0.8656 | 0.0347 | NS |
| 26: CPS II | 3.3310 | 0.1915 | NS | 0.6993 | 0.0402 | NS |
| 27: DEAN2 | 2.1038 | 0.3399 | NS | 0.4515 | 0.0723 | NS |
| 28: DEAN2 | 2.9155 | 0.6412 | NS | 0.5984 | 0.1334 | NS |
| 29: HUMBLE | 4.2780 | 0.2807 | 0.0091 | 0.9005 | 0.0590 | 0.0148 |
| 30: KAISE2 | 3.8758 | 0.4357 | NS | 0.8202 | 0.0922 | NS |
| 31: KAISE2 | 4.7883 | 0.5320 | NS | 1.0198 | 0.1130 | NS |
| 32: KATSOU | 2.9305 | 0.7339 | NS | 0.6031 | 0.1527 | NS |
| 33: LIAW | 2.4277 | 0.4372 | NS | 0.5093 | 0.0921 | NS |
| 34: MATOS | 3.4356 | 0.5336 | NS | 0.7276 | 0.1124 | NS |
| 35: PEZZO2 | 3.3022 | 0.5283 | 0.0002 | 0.7260 | 0.1137 | 0.0005 |

[1]Years smoked; [2]Not significant (NS) indicates $P \geq 0.1$. *P* values in the range $0.05 \leq P < 0.1$ are shown in brackets.

tors are significant at $P < 0.05$, including number of lung cancer cases ($P < 0.001$), continent ($P < 0.01$), publication year ($P < 0.01$), sex ($P < 0.05$) and product smoked ($P < 0.05$), though in a forward stepwise model (details not shown), only two factors were included: number of lung cancer cases ($P < 0.001$) and midpoint age ($P < 0.05$). However, the relationship of β1 to number of cases was not smooth (higher risks for 100 to < 200, and 500 to < 1000 cases and lower risks for < 100, 200 to < 500 and 1000+) so the result is difficult to interpret. The association with age is related to a lower β1 in older subjects (aged 60+ years).

For the 31 blocks combined, the RRs predicted for age of start 12.5, 15, 17.5, 20, 25 and 30 years are, respectively, 8.94, 7.80, 6.83, 5.99, 4.66 and 3.66 for the power model and 8.31, 7.57, 6.91, 6.30, 5.24 and 4.36 for the linear model (Figure 1D).

## DISCUSSION

In our earlier review[1] of the evidence relating smoking to lung cancer, we demonstrated a clear dose-response with the three measures considered, risk increasing with increasing amount smoked and duration and with decreasing age of starting. We extend this work by fitting parametric models to the dose-relationships.

We tried various models. The most useful were the "linear model" (log RR = $\beta_1 d$), the "power model" (log RR = $\beta_1 d^Y$) and the "log-with-baseline model" [log RR = $\beta_1 \log (1 + Wd)$], where d is dose. For amount smoked, the linear model proved inadequate, but a reasonable fit was found with the other two models, the best fit being for the log-with-baseline model with W = 0.81. For duration, all three models were reasonable, the best being the power model with Y = 0.74. For age of starting, where we used a duration-like dose measure based on (70 - age of starting to smoke), the best-fitting model was again the power model, here with Y = 1.44.

Inverse-variance weighted analyses were also carried out to identify sources of heterogeneity in β1. For amount smoked, as expected from our earlier review[1], the major source was continent, the fitted slope being much less

**Table 11 Duration of smoking by current smokers - inverse-variance weighted simple regression analyses of β1 based on best-fitting power model**

| Factor | Level | *n* | β1 (95%CI) | *P*[1] | RR for 30 yr smoked |
|---|---|---|---|---|---|
| All | | 35 | 0.79 (0.72-0.86) | | 5.94 |
| Sex | Male | 18 | 0.84 (0.73-0.94) | NS | 6.59 |
| | Female | 15 | 0.74 (0.64-0.85) | | 5.35 |
| | Combined | 2 | 0.79 (0.45-1.12) | | 5.89 |
| Study type | Case-control | 7 | 0.73 (0.50-0.97) | NS | 5.24 |
| | Prospective | 28 | 0.80 (0.72-0.87) | | 6.02 |
| Continent | North America | 27 | 0.81 (0.73-0.88) | NS | 6.19 |
| | Europe | 5 | 0.55 (0.20-0.90) | | 3.44 |
| | Asia | 1 | 0.51 (-0.11-1.13) | | 3.15 |
| | Other | 2 | 0.73 (0.19-1.27) | | 5.15 |
| Publication year[2] | < 1990 | 8 | 0.76 (0.52-0.99) | NS | 5.50 |
| | 1990-1994 | 2 | 0.44 (-0.20-1.09) | | 2.73 |
| | 1995-1999 | 25 | 0.80 (0.72-0.88) | | 6.08 |
| Product[3] | Any product | 4 | 0.50 (0.18-0.83) | NS | 3.12 |
| | Cigarettes +/- | 10 | 0.86 (0.73-0.98) | | 6.89 |
| | Cigarettes only | 21 | 0.78 (0.70-0.87) | | 5.85 |
| Unexposed | Never cigarettes | 19 | 0.77 (0.69-0.86) | NS | 5.71 |
| | Never any product | 16 | 0.85 (0.71-0.99) | | 6.75 |
| Grouped midpoint age (yr) | < 50 | 8 | 1.06 (0.78-1.34) | < 0.05 | 11.00 |
| | 50-59 | 13 | 0.87 (0.75-0.99) | | 7.04 |
| | 60+ | 14 | 0.73 (0.65-0.82) | | 5.21 |
| Cases in smokers | < 100 | 14 | 0.62 (0.39-0.84) | NS | 4.01 |
| | 100 to < 200 | 9 | 0.73 (0.59-0.87) | | 5.19 |
| | 200 to < 500 | 6 | 0.78 (0.67-0.89) | | 5.18 |
| | 500 to < 1000 | 5 | 0.95 (0.80-1.10) | | 8.58 |
| | 1000+ | 1 | 0.80 (0.59-1.01) | | 6.04 |

[1]Probability values for factor considered, presented as < 0.001, < 0.01, < 0.05, < 0.1 or not significant (NS) ($P \geqslant 0.1$); [2]Of principal publication for the study; [3]Any product = Smokes cigarettes and/or pipes and/or cigars; Cigarettes +/- = Smokes cigarettes with or without other products (pipes, cigars).

**Table 12 Age of starting to smoke by current smokers (yr) - dose-response data**

| Block: Study | Age of starting to smoke groupings (yr) | Mean values | RRs[1] |
|---|---|---|---|
| 1: CEDERL | < 17, 17-18, 19+ | 13.83, 17.52, 22.45 | 6.40, 9.80, 6.50 |
| 2: CEDERL | < 17, 17-18, 19+ | 14.31, 17.56, 24.32 | 0.61, 1.84, 1.99 |
| 3: CPS I | < 15, 15-19, 20-24, 25+ | 11.98, 16.74, 21.29, 28.47 | 15.00, 9.71, 7.14, 3.43 M |
| 4: CPS I | < 15, 15-19, 20-24, 25+ | 11.59, 16.56, 21.00, 30.90 | 18.16, 16.32, 12.00, 5.21 M |
| 5: CPS I | < 15, 15-19, 20-24, 25+ | 11.19, 16.61, 21.07, 35.00 | 14.03, 16.60, 8.66, 1.71 |
| 6: CPS I | < 15, 15-19, 20-24, 25+ | 12.71, 16.88, 21.26, 31.51 | 9.00, 5.00, 4.00, 1.50 M |
| 7: CPS I | < 20, 20-24, 25+ | 16.02, 20.95, 33.29 | 2.59, 3.23, 2.62 |
| 8: DEAN3 | < 15, 15-19, 20-24, 25+ | 11.73, 16.67, 21.19, 29.95 | 5.67, 7.01, 6.86, 6.80 |
| 9: DEAN3 | < 15, 15-19, 20-24, 25+ | 12.45, 16.95, 21.15, 32.19 | 2.41, 2.90, 2.95, 3.70 |
| 10: DORN | < 15, 15-19, 20-24, 25+ | 11.63, 16.43, 21.23, 30.88 | 23.42, 16.25, 11.06, 5.18 M |
| 11: DORN | < 15, 15-19, 20-24, 25+ | 11.37, 16.81, 20.69, 32.63 | 14.18, 11.31, 7.94, 4.95 M |
| 12: ENGELA | < 20, 20-29, 30+ | 15.08, 22.23, 35.10 | 7.42, 3.60, 3.79 |
| 13: ENGELA | < 20, 20-29, 30+ | 15.80, 22.63, 35,80 | 11.29, 8.15, 2.73 M |
| 14: GAO2 | < 20, 20-29, 30+ | 15.11, 22.17, 35.94 | 8.62, 6.44, 2.15 M |
| 15: HIRAYA | < 20, 20+ | 14.90, 24.02 | 5.71, 4.35 M |
| 16: HIRAYA | < 20, 20+ | 15.87, 27.40 | 0.78, 2.46 |
| 17: LIAW | < 21, 21-24, 25+ | 15.70, 21.94, 32.03 | 4.60, 5.90, 1.50 |
| 18: MATOS | < 15, 15-19, 20+ | 11.96, 16.64, 23.49 | 11.30, 8.60, 5.30 M |
| 19: MIGRAN | < 16, 16-19, 20+ | 12.60, 17.19, 23.71 | 10.03, 6.93, 7.79 |
| 20: MIGRAN | < 16, 16-19, 20+ | 12.94, 17.31, 26.17 | 7.17, 8.29, 7.98 |
| 21: MRFITR | < 16, 16-17, 18-19, 20-21, 22-23, 24+ | 12.93, 16.45, 18.31, 20.40, 22.35, 27.74 | 45.91, 67.17, 50.54, 27.09, 60.06, 23.91 |
| 22: SEGI2 | < 20, 20-22, 23+ | 15.28, 20.77, 26.60 | 8.21, 5.56, 1.83 M |
| 23: SEGI2 | < 20, 20-22, 23+ | 14.87, 20.90, 28.96 | 8.70, 5.68, 3.56 M |
| 24: SEGI2 | < 20, 20-22, 23+ | 14.37, 20.59, 30.87 | 3.26, 1.70, 1.52 M |
| 25: SVENSS | < 19, 19-25, 26+ | 15.04, 21.31, 33.89 | 7.82, 13.08, 5.61 |
| 26: WAKAI | < 20, 20-29, 30+ | 14.91, 22.14, 37.23 | 3.69, 4.62, 2.08 |
| 27: WU | < 19, 19-24, 25+ | 15.06, 20.71, 31.44 | 10.32, 3.57, 1.55 M |

[1]M indicates a strictly monotonic decline in RR with increasing age of starting to smoke.

**Table 13 Comparing the suitability of different models relating log *RR* to age of starting smoke by current smokers, expressed as d = (70-age at start)/10**

| Model | Parameter value[1] | Fitted coefficient(s) (SE) | Deviance | DF | Deviance explained (%) |
|---|---|---|---|---|---|
| Null | - | - | 2145.30 | 27 | |
| Linear: log RR = $\beta_1 d$ | - | $\beta_1$ = 0.3681 (0.0085) | 276.67 | 26 | 87.10 |
| Quadratic: log RR = $\beta_1 d + \beta_2 d^2$ | - | $\beta_1$ = 0.1987 (0.0349)<br>$\beta_2$ = 0.0318 (0.0064) | 251.63 | 25 | 88.27 |
| Cubic: log RR = $\beta_1 d + \beta_2 d^2 + \beta_3 d^3$ | - | $\beta_1$ = -0.0415 (0.2143)<br>$\beta_2$ = 0.1304 (0.0870)<br>$\beta_3$ = -0.0100 (0.0088) | 250.34 | 24 | 88.33 |
| Power: log RR = $\beta_1 d^Y$ | Y = 0.75 | $\beta_1$ = 0.5515 (0.0129) | 316.69 | 26 | |
| | Y = 1.00 | $\beta_1$ = 0.3681 (0.0085) | 276.67 | 26 | |
| | Y = 1.42 | $\beta_1$ = 0.1825 (0.0042) | 251.27 | 26 | |
| | Y = 1.43 | $\beta_1$ = 0.1794 (0.0041) | 251.24 | 26 | |
| | Y = 1.44 | $\beta_1$ = 0.1764 (0.0041) | 251.23 | 26 | 88.29 |
| | Y = 1.45 | $\beta_1$ = 0.1734 (0.0040) | 251.25 | 26 | |
| | Y = 1.46 | $\beta_1$ = 0.1705 (0.0039) | 251.28 | 26 | |
| | Y = 1.50 | $\beta_1$ = 0.1592 (0.0037) | 251.67 | 26 | |
| | Y = 2.00 | $\beta_1$ = 0.0668 (0.0015) | 284.04 | 26 | |
| Log: log RR = $\beta_1 \log d$ | - | $\beta_1$ = 1.1460 (0.0270) | 340.23 | 26 | 84.14 |
| Exponential: log RR = $\beta_1 \exp d$ | - | $\beta_1$ = 0.0058 (0.0002) | 852.85 | 26 | 60.24 |

[1]Note that we only sought the best-fitting value of Y to two decimal places.

**Table 14 Age of starting to smoke by current smokers - observed and fitted lung cancers for the linear and best power model, with $\beta_1$ fitted separately for each block**

| Age of starting (yr) | Observed[1] | Fitted[2] | |
|---|---|---|---|
| | | Linear model | Best power model |
| < 16 | 1304.92 | 1294.48 | 1387.73 |
| 16 to < 20 | 1964.76 | 1906.55 | 1921.20 |
| 20 to < 24 | 1227.48 | 1266.99 | 1216.48 |
| 25+ | 2173.90 | 2219.26 | 2122.14 |
| Never smoked | 894.41 | 878.20 | 917.92 |
| Total | 7565.47 | 7565.47 | 7565.47 |
| Fit statistic[3] | | 4.77 | 9.75 |

[1]Observed pseudo-number of lung cancer cases, summed over blocks; [2]Fitted pseudo-number of lung cancer cases, summed over blocks. For each study, the fitted number of cases for each block is calculated from the numbers at risk and the fitted relative risks for age of starting to smoke, derived from the fitted value of $\beta_1$; [3]Based on summation of (observed-fitted)$^2$/fitted, the summation also including terms for the observed and fitted total numbers of controls (not shown). The statistic can be considered to be approximately chisquared distributed on 6 df and is not significant ($P > 0.1$) for both models.

steep for Asian than European or North American studies. However, it proved more difficult to identify meaningful major sources for the other measures.

We discuss various issues relating to interpretation of these findings.

### Adequacy of literature search and publication bias

All the data used came from the IESLC database. The source paper[1] demonstrated that the search was comprehensive, though limited to papers published before 2000 and studies of 100+ cases. Publication bias was discussed earlier[1], evidence for its existence being considered not strong. The probability of dose-response results being published might depend on the strength of the overall relationship seen. While clearly demonstrated for passive smoking and lung cancer[3], this seems less relevant here, the association with active smoking being so strong. Nevertheless, some publication bias may exist.

There are various reasons why the fitted dose-relationships may not accurately reflect the true relationships.

### Misclassification of smoking status

It is well-documented (*e.g.*,[12,13]) that some subjects deny current or past smoking, so increasing the apparent lung cancer risk in reported never smokers and biasing downwards the estimated smoking RR. Such misclassification is difficult to adjust for, as it varies by aspects of study design, the questions asked, and also by sex, age, location and other demographics. Indeed, higher denial rates in Asian populations[14] may contribute to the markedly weaker observed associations seen in Asia.

In prospective studies there is an additional problem, especially in studies with long-term follow-up with no re-interviews to update smoking status. In particular, some subjects classified at baseline as current smokers may quit during follow-up. Also some never smokers may start, though this is less likely given the subjects' age at baseline in many studies.

### Misclassification of amount smoked

Similar problems arise. Subjects may understate (or overstate) the amount they smoke, and during follow-up in prospective studies, may reduce or increase the amount smoked. Although some studies, particularly case-control, may ask questions on habits at various times during the subject's smoking career, the data reported may relate to average consumption. Someone smoking, say, 30 cigarettes/d for 20 years, then 10 cigarettes/d for 20 years, may not have the same risk as someone smoking 20 cigarettes/d for the whole 40 years period. Difficulties in re-

**Table 15 Age of starting to smoke by current smokers - fitted values of $\beta_1$ and SE, and *P* values for goodness-of-fit tests for the linear model and for the best-fitting power model**

| Block: Study | Linear model[1] log RR = $\beta_1$(70 - a)/10 | | | Power model[1] log RR = $\beta_1$(70 - a)/$10^{1.44}$ | | |
|---|---|---|---|---|---|---|
| | $\beta_1$ | SE $\beta_1$ | *P* (fit)[2] | $\beta_1$ | SE $\beta_1$ | *P* (fit)[2] |
| 1: CEDERL | 0.0385 | 0.0082 | NS | 0.1828 | 0.0394 | NS |
| 2: CEDERL | 0.0140 | 0.0089 | NS | 0.0699 | 0.0445 | NS |
| 3: CPS I | 0.0461 | 0.0041 | 0.0490 | 0.2167 | 0.0188 | NS |
| 4: CPS I | 0.0524 | 0.0029 | (0.0959) | 0.2399 | 0.0135 | (0.0579) |
| 5: CPS I | 0.0512 | 0.0047 | 0.0195 | 0.2408 | 0.0223 | 0.0307 |
| 6: CPS I | 0.0301 | 0.0036 | (0.0510) | 0.1480 | 0.0175 | NS |
| 7: CPS I | 0.0237 | 0.0046 | NS | 0.1168 | 0.0238 | 0.0828 |
| 8: DEAN3 | 0.0332 | 0.0041 | NS | 0.1482 | 0.0189 | 0.0342 |
| 9: DEAN3 | 0.0209 | 0.0039 | NS | 0.0960 | 0.0187 | NS |
| 10: DORN | 0.0549 | 0.0037 | (0.0874) | 0.2527 | 0.0166 | NS |
| 11: DORN | 0.0455 | 0.0027 | NS | 0.2086 | 0.0125 | NS |
| 12: ENGELA | 0.0360 | 0.0037 | 0.0109 | 0.1682 | 0.0171 | 0.0269 |
| 13: ENGELA | 0.0442 | 0.0045 | NS | 0.2186 | 0.0221 | NS |
| 14: GAO2 | 0.0394 | 0.0066 | NS | 0.1909 | 0.0318 | NS |
| 15: HIRAYA | 0.0317 | 0.0022 | NS | 0.1497 | 0.0104 | 0.0288 |
| 16: HIRAYA | 0.0207 | 0.0029 | NS | 0.1086 | 0.0152 | NS |
| 17: LIAW | 0.0296 | 0.0050 | 0.0396 | 0.1448 | 0.0242 | (0.0661) |
| 18: MATOS | 0.0411 | 0.0063 | NS | 0.1933 | 0.0296 | NS |
| 19: MIGRAN | 0.0373 | 0.0087 | NS | 0.1612 | 0.0385 | NS |
| 20: MIGRAN | 0.0400 | 0.0108 | NS | 0.1859 | 0.0519 | NS |
| 21: MRFITR | 0.0686 | 0.0237 | NS | 0.2931 | 0.1055 | NS |
| 22: SEGI2 | 0.0468 | 0.0140 | NS | 0.2325 | 0.0639 | NS |
| 23: SEGI2 | 0.0440 | 0.0128 | NS | 0.1974 | 0.0557 | NS |
| 24: SEGI2 | 0.0192 | 0.0109 | NS | 0.0945 | 0.0500 | NS |
| 25: SVENSS | 0.0432 | 0.0051 | NS | 0.2064 | 0.0246 | NS |
| 26: WAKAI | 0.0275 | 0.0068 | NS | 0.1244 | 0.0318 | NS |
| 27: WU | 0.0348 | 0.0063 | NS | 0.1741 | 0.0305 | NS |

[1]a = Age of starting to smoke; [2]Not significant (NS) indicates $P \geqslant 0.1$. *P* values in the range $0.05 \leqslant P < 0.1$ are shown in brackets.

**Table 16 Age of starting to smoke by current smokers - inverse-variance weighted simple regression analyses of $\beta_1$ based on best-fitting power model**

| Factor | Level | *n* | $\beta_1$ (95%CI) | *P*[1] | RR for 15 yr start |
|---|---|---|---|---|---|
| All | | 27 | 0.18 (0.16-0.20) | | 7.80 |
| Sex | Male | 17 | 0.19 (0.17-0.21) | < 0.05 | 9.49 |
| | Female | 9 | 0.14 (0.11-0.17) | | 5.14 |
| | Combined | 1 | 0.14 (0.04-0.25) | | 5.40 |
| Study type | Case-control | 10 | 0.15 (0.11-0.20) | NS | 5.94 |
| | Prospective | 17 | 0.18 (0.16-0.20) | | 8.37 |
| Continent | North America | 9 | 0.21 (0.19-0.23) | < 0.01 | 11.49 |
| | Europe | 9 | 0.16 (0.13-0.19) | | 6.38 |
| | Asia | 8 | 0.14 (0.11-0.17) | | 5.17 |
| | Other | 1 | 0.19 (0.08-0.31) | | 9.50 |
| Publication year[2] | < 1990 | 11 | 0.15 (0.11-0.19) | < 0.01 | 5.59 |
| | 1990-1994 | 4 | 0.14 (0.11-0.17) | | 5.17 |
| | 1995-1999 | 12 | 0.20 (0.18-0.22) | | 10.37 |
| Product[3] | Any product | 3 | 0.17 (0.11-0.24) | < 0.05 | 7.65 |
| | Cigarettes +/- | 16 | 0.19 (0.17-0.21) | | 9.09 |
| | Cigarettes only | 8 | 0.13 (0.09-0.17) | | 4.69 |
| Unexposed | Never cigarettes | 4 | 0.19 (0.13-0.25) | NS | 9.05 |
| | Never any product | 23 | 0.17 (0.15-0.20) | | 7.66 |
| Grouped midpoint age (yr) | < 50 | 7 | 0.18 (0.13-0.23) | NS | 7.73 |
| | 50-59 | 9 | 0.21 (0.17-0.25) | | 10.98 |
| | 60+ | 11 | 0.17 (0.14-0.19) | | 6.88 |
| Cases in smokers | < 100 | 12 | 0.14 (0.11-0.17) | < 0.001 | 5.09 |
| | 100 to < 200 | 7 | 0.20 (0.16-0.24) | | 9.88 |
| | 200 to < 500 | 4 | 0.17 (0.23-0.20) | | 7.02 |
| | 500 to < 1000 | 3 | 0.23 (0.20-0.26) | | 14.57 |
| | 1000+ | 1 | 0.15 (0.11-0.19) | | 5.71 |

[1]Probability values for factor considered, presented as < 0.001, < 0.01, < 0.05, < 0.1 or not significant (NS) ($P \geqslant 0.1$); [2]Of principal publication for the study; [3]Any product = Smokes cigarettes and/or pipes and/or cigars; Cigarettes +/- = Smokes cigarettes with or without other products (pipes, cigars).

membering smoking history also form part of the problem. Also the dose of smoke constituents received may not be directly proportional to the amount smoked[15].

### Misclassification of duration and age of starting

Subjects may not remember the exact age of starting, and indeed there may be differences in definition between studies - age of first trying a cigarette, or age of starting to smoke regularly? Also duration may not represent a continuous period. Risk may be affected by intermediate quit periods, which may be asked about differently in different studies.

### Estimating midpoints of ranges

The statistical methods used require estimates of midpoints of ranges used. We have not attempted sensitivity analyses based on alternative procedures for defining midpoints.

### Use of pseudo-numbers

Our methodology requires knowledge, for each block, of the numbers of cases and controls (or at risk) in each smoking group. As such data are not always provided, and indeed for covariate-adjusted data are only hypothetical, we used the method of Hamling *et al*[4] to estimate pseudo-numbers corresponding exactly to the reported RRs and CIs. These pseudo-numbers have been shown[16] to allow accurate estimation of RRs and CIs relative to a different base group from that used originally, and should be adequate for model fitting. This issue seems less important than others considered so far.

### Adjustment for other smoking variables

Our analyses compare risk relative to never smokers, all the RRs in any block being adjusted for the same variables. As RRs relative to never smokers cannot be adjusted for other smoking variables, we necessarily restricted attention to estimates adjusted for age and non-smoking characteristics. This is possibly unfortunate as, for example, later starters may smoke less than earlier starters. In theory one could study the extent of such bias based on studies presenting risk (compared to never smokers) jointly by more than one dose measure. However, few studies present such data and we did not investigate this.

### Use of simple models based on published results

We restricted attention to models of a relatively simple functional form, partly as it is much easier to explain results and conduct tests of heterogeneity where differences between blocks can be expressed in terms of one parameter ($\beta_1$). Also, the numerous data uncertainties may not justify a more complex approach. Such an approach is better pursued using individual person data from large studies. This would allow fitting of models simultaneously accounting for amount smoked and duration, and allow a more precise risk estimation. In the context of a systematic review and meta-analysis, involving many studies conducted years ago with the data unlikely to be accessible, we made no attempt to obtain individual data sets.

### Model fit

Goodness-of-fit has been studied in various ways. First, we used the "pool-first" approach[9,10] to compare the deviance of models with a common $\beta_1$ per block but a different functional form of the dose-relationship. The exponential (log RR = $\beta_1$exp d) and the log model (log RR = $\beta_1$log d) clearly fitted substantially worse than other models, and were not pursued further. Also, the power model (log RR = $\beta_1 d^Y$) and the log-with-baseline model [log RR = $\beta_1$log (1 + Wd)] generally fitted better than the linear model (log RR = $\beta_1$d), though for age of starting the best-fitting log-with-baseline model had such a low estimate of W that it became equivalent to the linear model. While the deviance of the linear model was reduced by adding quadratic and cubic terms this advantage was small. We concentrated most on the power and/or log-with-baseline models, given their greater simplicity, and the fact that the cubic model fitted worse than these alternatives for amount smoked and not materially better for duration or age of start.

We then restricted attention to the linear, power and, except for age of start, the log-with-baseline model, fitting separate $\beta_1$ values to each block. We investigated goodness-of-fit by studying plots of observed and predicted RRs (not shown), and by comparing observed and predicted numbers, both within block (Fit Amount Smoked[11], Fit Duration, and Fit Age Start[11]) and summed over block (Tables 4, 9 and 14). This allowed two general conclusions. First, the best models (log-with-baseline for amount smoked, power for duration and age of start) fitted the shape of the dose relationship well. Given the large number of cases analyzed (54245 for amount smoked, 10711 for duration and 7575 for age of start) it is unsurprising that formal misfit existed for amount smoked and duration, but this seems relatively unimportant. Second, there were significant misfits for some blocks. The results section comments on the worst cases. Sometimes these are due to unusual response patterns, difficult to fit by any plausible model, sometimes to differing response patterns in different blocks. Thus, for amount smoked, there are some blocks where the slope flattens off at high consumption, but others where the reverse is true. The explanation for this is unclear, but attempting to account for it by more complex models seems unattractive, as compared to the models selected, which involve a common shape and variation only in slope ($\beta_1$).

### Sources of heterogeneity

We carried out weighted regression analyses to investigate sources of heterogeneity. While some factors (*e.g.*, age and sex) could be better evaluated using pooled analyses based on individual person data, and problems arise from correlations between variables studied, these analyses should detect major sources.

For amount smoked, these analyses only identified

continent as a significant factor, other associations seen in the simple analyses becoming non-significant once continent was accounted for. The smaller $\beta_1$ for Asian studies is consistent with our earlier analyses[16], and may relate to higher denial rates of smoking in Asia.

For duration and age of starting, the regression analyses showed a tendency for $\beta_1$ to be greater in studies involving more lung cancer cases and studies of younger people. Higher values in males than females and lower values in Asian studies were not independently significant. There was also some evidence for duration of higher $\beta_1$ values for smoking cigarettes, than smoking any product.

### Comparison with some previous work

Attempts have been made before to model the relationship of lung cancer to amount smoked and duration. For example, Doll *et al*[17], in a much cited paper, based on data for British doctors who started smoking at ages 16-25 and smoked 40 or less per day, modelled the annual lung cancer incidence at age 40-79 by the expression

$0.273 \times 10^{-12} \times (\text{cigarettes/d} + 6)^2 \times (\text{age} - 22.5)^{4.5}$.

They noted "significant ($P < 0.01$) upward curvature of the dose-response relationship in the range 0-40 cigarettes/d, which is what might be expected if more than one of the 'stages' (in the multistage genesis of bronchial carcinoma) was strongly affected by smoking." They also noted a drop off in response above 40 cigarettes/d, though based on few cases, and discussed various explanations for it. Our analyses show little evidence of upward curvature with amount smoked. However, this does not rule out smoking affecting more than one stage of a multistage process; indeed there is strong evidence this is true[18].

Taking (age - 22.5 years) as an approximate indicator of duration, the model of Doll *et al*[17] suggests risk rises steeply with increasing duration, according to a fourth or fifth power relationship. At first sight, this appears to conflict with our findings, where the power relationship we fitted was only somewhat above linear (Figure 1C). However, whereas Doll and Peto's analysis compares risk by age for people of a similar age of start, our modelling compares risk by age of start for people of a given age. Here, the relationship of risk to duration will be much less steep. This can be illustrated by applying formulae for a form of the multistage model where risk affects the first and penultimate stages, the effect on the penultimate stage being twice as strong as for the first stage, a form known to fit smoking and lung cancer relationships quite well[18,19]. The RR for a 70-year-old starting at age 15 is estimated as 1.66 times higher than for a 70-year-old starting at age 30. This ratio somewhat exceeds the ratio of durations (55/40 = 1.38), but much less than predicted by a fourth or fifth power relationship ($1.38^{4.5} = 4.26$).

### Summing up

Based on 71 studies described in 87 publications[20-106] we demonstrated that for all three measures of dose studied (amount smoked, duration and age of start), the shape of their relationship with lung cancer can be described quite accurately using simple models. Though, for all dose measures, there is evidence of misfit for some data blocks, these seem mainly due to unusual response patterns difficult to fit with plausible models, or to different blocks showing differing shapes of the dose-relationship. The main limitations of the models relate to the data they were fitted to. Misclassification of smoking status and of dose may produce bias, as may failure to update smoking habits during follow-up in prospective studies, and failure to adjust for other indices of dose. Nevertheless, the models presented characterize the observed relationships of lung cancer to amount smoked, duration, and age of start more fully than previously attempted.

## ACKNOWLEDGMENTS

We thank Philip Morris Products S.A. for supporting this research. The opinions and conclusions of the authors are their own, and do not necessarily reflect the position of Philip Morris Products S.A. We also thank Pauline Wassell, Diana Morris and Yvonne Cooper for assistance in typing various drafts of the paper and obtaining relevant literature.

## COMMENTS

### Background

No previous meta-analysis has used parametric models in order to quantify more precisely the relationship between smoking and lung cancer. Using a database of all epidemiological studies of 100 or more lung cancer cases published before 2000, models are fitted relating lung cancer risk to amount smoked, duration of smoking and age of starting to smoke.

### Research frontiers

Based on all the studies providing relevant data, the models fitted show that the risk, relative to never smokers, rises from 3.86 to 22.31 between 10 and 50 cigarettes/d, from 2.21 to 13.54 between 10 and 50 years smoked, and from 3.66 to 8.94 between age of starting 30 and 12.5 years. There is little heterogeneity between studies for duration of smoking or age started, but there is clear heterogeneity for amount smoked, with RRs for 50 cigarettes/d being 7.23 for studies in Asia, as compared to 26.36 for North American and 22.16 for European studies.

### Innovations and breakthroughs

The new feature of this paper is the comprehensive assessment of the shape of the dose-responses studied, with a number of alternative functional forms studied, and the best-fitting one selected. The fitted models, which describe the relationships well, are each quite simple in form, allowing ready meta-analysis of individual study estimates.

### Applications

The fitted models allow more precise quantification of the hazards of smoking than previously reported, and will assist smoking and health researchers.

### Terminology

Linear model: The logarithm of the RR is linearly related to dose. In the power model it is related to dose raised to a power. In the log-with-baseline model, it is related to the logarithm of dose with an offset for background risk. Pseudo-numbers are estimates of numbers of cases and controls, by dose level, derived from published RRs, which allow fitting of the models.

### Peer review

The authors meta-analyzed smoking/lung cancer relationships using parametric modelling according to the IESLC database. They found that the models describe the dose-relationship well and concluded that they can be used to more precisely estimate the lung cancer risk from smoking. The limitation has been fully discussed in the discussion part. This study provides some interesting results for further research into smoking and lung cancer.

## REFERENCES


1 **Lee PN**, Forey BA, Coombs KJ. Systematic review with meta-analysis of the epidemiological evidence in the 1900s relating smoking to lung cancer. *BMC Cancer* 2012; **12**: 385 [PMID: 22943444 DOI: 10.1186/1471-2407-12-385]

2 **Fry JS**, Lee PN, Forey BA, Coombs KJ. How rapidly does the excess risk of lung cancer decline following quitting smoking? A quantitative review using the negative exponential model. *Regul Toxicol Pharmacol* 2013; **67**: 13-26 [PMID: 23764305 DOI: 10.1016/j.yrtph.2013.06.001]

3 **Fry JS**, Lee PN. Revisiting the association between environmental tobacco smoke exposure and lung cancer risk. I. The dose-response relationship with amount and duration of smoking by the husband. *Indoor Built Environ* 2000; **9**: 303-316 [DOI: 10.1177/1420326X0000900602]

4 **Hamling J**, Lee P, Weitkunat R, Ambühl M. Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category. *Stat Med* 2008; **27**: 954-970 [PMID: 17676579 DOI: 10.1002/sim.3013]

5 Thun MJ, Day-Lally C, Myers DG, Calle EE, Flanders WD, Zhu BP, Namboodiri MM, Heath CW Jr. Trends in tobacco smoking and mortality from cigarette use in cancer prevention studies I (1959 through 1965) and II (1982 through 1988). In: Shopland DR, Burns DM, Garfinkel L, Samet JM, editors. Changes in cigarette-related disease risks and their implications for prevention and control. Rockville, Maryland: US Department of Health and Human Services, National Institutes of Health, National Cancer Institute, 1997: 305-382. Available from: URL: http://cancercontrol.cancer.gov/tcrb/monographs/8/m8_4.pdf

6 **Forey B**, Hamling J, Lee P, Wald N. International Smoking Statistics: A collection of historical data from 30 economically developed countries. 2nd ed. London and Oxford: Wolfson Institute of Preventive Medicine and Oxford University Press, 2002 [DOI: 10.1093/acprof:oso/9780198508564.001.0001]

7 **Forey B**, Hamling J, Lee P. International Smoking Statistics. A collection of worldwide historical data. 2nd ed. Sutton: P N Lee Statistics and Computing Ltd., 2006-2012. Available from: URL: http://www.pnlee.co.uk/iss.htm

8 **US Department of Health and Human Services**. National health and nutrition examination survey (NHANES). National Center for Health Statistics. Available from: URL: http://www.cdc.gov/nchs/nhanes.htm

9 **Greenland S**, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol* 1992; **135**: 1301-1309 [PMID: 1626547]

10 **Berlin JA**, Longnecker MP, Greenland S. Meta-analysis of epidemiologic dose-response data. *Epidemiology* 1993; **4**: 218-228 [PMID: 8512986 DOI: 10.1097/00001648-199305000-00005]

11 **http://www.pnlee.co.uk/downloads/ieslc3/lc_dose_response_paper_additionalfiles.htm**

12 **Lee PN**. Misclassification of smoking habits and passive smoking. A review of the evidence. Heidelberg: Springer-Verlag, 1988 [DOI: 10.1007/978-3-642-73822-7]

13 **Lee PN**, Forey BA. Misclassification of smoking habits as determined by cotinine or by repeated self-report - a summary of evidence from 42 studies. *J Smoking-Related Dis* 1995; **6**: 109-129

14 **Lee PN**, Forey B, Fry JS. Revisiting the association between environmental tobacco smoke exposure and lung cancer risk. III. Adjustment for the biasing effect of misclassification of smoking habits. *Indoor Built Environ* 2001; **10**: 384-398 [DOI: 10.1177/1420326X0101000605]

15 **Baker RR**, Dixon M, Mariner DC, Shepperd CJ, Scherer G, Ogden MW, Robinson JH, Sinclair NM, Sherwood N, Akiyama Y, Sakamoto K, Röper W, Tricker AR, Marchand V, Varignon B, Lionetti G. Terms used for exposure to smoke. *Beitr Tabakforsch Int* 2004; **21**: 250

16 **Orsini N**, Li R, Wolk A, Khudyakov P, Spiegelman D. Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software. *Am J Epidemiol* 2012; **175**: 66-73 [PMID: 22135359]

17 **Doll R**, Peto R. Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and lifelong non-smokers. *J Epidemiol Community Health* 1978; **32**: 303-313 [PMID: 744822 DOI: 10.1136/jech.32.4.303]

18 **Lee PN**. Studying the relationship of smoking to lung cancer using the multistage model of carcinogenesis. A review. Sutton, Surrey: P N Lee Statistics and Computing Ltd., 1995. Available from: URL: http://www.pnlee.co.uk/Reports.htm

19 **Day NE**, Brown CC. Multistage models and primary prevention of cancer. *J Natl Cancer Inst* 1980; **64**: 977-989 [PMID: 6929006]

20 **Akiba S**. Analysis of cancer risk related to longitudinal information on smoking habits. *Environ Health Perspect* 1994; **102** Suppl 8: 15-19 [PMID: 7851325 DOI: 10.1289/ehp.94102s815]

21 **Amandus H**, Costello J. Silicosis and lung cancer in U.S. metal miners. *Arch Environ Health* 1991; **46**: 82-89 [PMID: 2006898 DOI: 10.1080/00039896.1991.9937433]

22 **Archer VE**, Gillam JD, Wagoner JK. Respiratory disease mortality among uranium miners. *Ann NY Acad Sci* 1976; **271**: 280-293 [DOI: 10.1111/j.1749-6632.1976.tb23123.x]

23 **Rylander R**, Axelsson G, Andersson L, Liljequist T, Bergman B. Lung cancer, smoking and diet among Swedish men. *Lung Cancer* 1996; **14** Suppl 1: S75-S83 [PMID: 8785669 DOI: 10.1016/S0169-5002(96)90212-3]

24 **Modigh C**, Axelsson G, Alavanja M, Andersson L, Rylander R. Pet birds and risk of lung cancer in Sweden: a case-control study. *BMJ* 1996; **313**: 1236-1238 [PMID: 8939112 DOI: 10.1136/bmj.313.7067.1236]

25 **Ben-Shlomo Y**, Smith GD, Shipley MJ, Marmot MG. What determines mortality risk in male former cigarette smokers? *Am J Public Health* 1994; **84**: 1235-1242 [PMID: 8059878 DOI: 10.2105/AJPH.84.8.1235]

26 **Marmot MG**, Shipley MJ, Rose G. Inequalities in death--specific explanations of a general pattern? *Lancet* 1984; **1**: 1003-1006 [PMID: 6143919]

27 **Department of National Health and Welfare Canada**. A Canadian study of smoking and health. Canada: Department of National Health and Welfare, 1966

28 **Boucot KR**, Weiss W, Seidman H, Carnahan WJ, Cooper DA. The Philadelphia pulmonary neoplasm research project: basic risk factors of lung cancer in older men. *Am J Epidemiol* 1972; **95**: 4-16 [PMID: 5007365]

29 **Brett GZ**, Benjamin B. Smoking habits of men employed in industry, and mortality. *Br Med J* 1968; **3**: 82-85 [PMID: 5662965 DOI: 10.1136/bmj.3.5610.82]

30 **Bross ID**, Gibson R. Risks of lung cancer in smokers who switch to filter cigarettes. *Am J Public Health Nations Health* 1968; **58**: 1396-1403 [PMID: 5691372 DOI: 10.2105/AJPH.58.8.1396]

31 **Buffler PA**, Pickle LW, Mason TJ, Contant C. The causes of lung cancer in Texas. In: Mizell M, Correa P, editors. Lung cancer: causes and prevention, Proceedings of the International Lung Cancer Update Conference, New Orleans, Louisiana, March 3-5, 1983. Deerfield Beach, FL: Verlag Chemie International Inc., 1984: 83-99

32 **Ives JC**. Environmental exposures and lung cancer risk among women in Harris County, Texas, 1977-1980 [Thesis]. Houston, TX: University of Texas, Health Science Centre, 1984

33 **Cederlöf R**, Friberg L, Hrubec Z, Lorich U. The relationship of smoking and some social covariables to mortality and cancer morbidity. A ten year follow-up in a probability sample of 55,000 Swedish subjects age 18-69, Part 1/2. Stockholm:

Karolinska Institute, Dept of Environmental Hygiene, 1975
34 **Nordlund LA**, Carstensen JM, Pershagen G. Cancer incidence in female smokers: a 26-year follow-up. *Int J Cancer* 1997; **73**: 625-628 [PMID: 9398036]
35 **Chang AK**, Barrett-Connor E, Edelstein S. Low plasma cholesterol predicts an increased risk of lung cancer in elderly women. *Prev Med* 1995; **24**: 557-562 [PMID: 8610078 DOI: 10.1006/pmed.1995.1089]
36 **Chow WH**, Schuman LM, McLaughlin JK, Bjelke E, Gridley G, Wacholder S, Chien HT, Blot WJ. A cohort study of tobacco use, diet, occupation, and lung cancer mortality. *Cancer Causes Control* 1992; **3**: 247-254 [PMID: 1610971 DOI: 10.1007/BF00124258]
37 **Comstock GW**, Alberg AJ, Huang HY, Wu K, Burke AE, Hoffman SC, Norkus EP, Gross M, Cutler RG, Morris JS, Spate VL, Helzlsouer KJ. The risk of developing lung cancer associated with antioxidants in the blood: ascorbic acid, carotenoids, alpha-tocopherol, selenium, and total peroxyl radical absorbing capacity. *Cancer Epidemiol Biomarkers Prev* 1997; **6**: 907-916 [PMID: 9367064]
38 **Correa P**, Pickle LW, Fontham E, Dalager N, Lin Y, Haenszel W, Johnson WD. The causes of lung cancer in Louisiana. In: Mizell M, Correa P, editors. Lung cancer: causes and prevention. New York: Verlag Chemie International Inc., 1984: 73-82
39 **Fontham ET**, Pickle LW, Haenszel W, Correa P, Lin YP, Falk RT. Dietary vitamins A and C and lung cancer risk in Louisiana. *Cancer* 1988; **62**: 2267-2273 [PMID: 3179940]
40 **Burns DM**, Shanks TG, Choi W, Thun MJ, Heath CW Jr, Garfinkel L. The American Cancer Society cancer prevention study I: 12-year follow-up of 1 million men and women. In: Changes in cigarette-related disease risks and their implications for prevention and control. Rockville, MD: US Department of Health and Human Services, National Institutes of Health, National Cancer Institute, 1997: 113-304. Available from: URL: http://cancercontrol.cancer.gov/tcrb/monographs/8/m8_3.pdf
41 **Hammond EC**. Smoking habits and air pollution in relation to lung cancer. In: Lee HK, editor. Environmental factors in respiratory diseases. New York: Academic Press Inc., 1972: 177-198
42 **Thun MJ**, Myers DG, Day-Lally C, Namboodiri MM, Calle EE, Flanders WD, Adams SL, Heath CW Jr. Age and the exposure-response relationships between cigarette smoking and premature death in Cancer Prevention Study II. In: Changes in cigarette-related disease risks and their implications for prevention and control. Rockville, MD: US Department of Health and Human Services, National Institutes of Health, National Cancer Institute, 1997: 383-475. Available from: URL: http://cancercontrol.cancer.gov/tcrb/monographs/8/m8_5.pdf
43 **Darby S**, Whitley E, Silcocks P, Thakrar B, Green M, Lomas P, Miles J, Reeves G, Fearn T, Doll R. Risk of lung cancer associated with residential radon exposure in south-west England: a case-control study. *Br J Cancer* 1998; **78**: 394-408 [PMID: 9703290 DOI: 10.1038/bjc.1998.506]
44 **Wicken AJ**. Environmental and personal factors in lung cancer and bronchitis mortality in Northern Ireland, 1960-1962. London: Tobacco Research Council, 1966
45 **Dean G**, Lee PN, Todd GF, Wicken AJ. Report on a second retrospective mortality study in North-East England - Part I. Factors related to mortality from lung cancer, bronchitis, heart disease and stroke in Cleveland County, with particular emphasis on the relative risks associated with smoking filter and plain cigarettes. London: Tobacco Research Council, 1977
46 **de Klerk NH**, Musk AW. Silica, compensated silicosis, and lung cancer in Western Australian goldminers. *Occup Environ Med* 1998; **55**: 243-248 [PMID: 9624278 DOI: 10.1136/oem.55.4.243]
47 **Doll R**, Peto R, Wheatley K, Gray R, Sutherland I. Mortality in relation to smoking: 40 years' observations on male British doctors. *BMJ* 1994; **309**: 901-911 [PMID: 7755693 DOI: 10.1136/bmj.309.6959.901]
48 **Doll R**, Peto R. Mortality in relation to smoking: 20 years' observations on male British doctors. *Br Med J* 1976; **2**: 1525-1536 [PMID: 1009386 DOI: 10.1136/bmj.2.6051.1525]
49 **Doll R**, Gray R, Hafner B, Peto R. Mortality in relation to smoking: 22 years' observations on female British doctors. *Br Med J* 1980; **280**: 967-971 [PMID: 7417764 DOI: 10.1136/bmj.280.6219.967]
50 **Dorant E**, van den Brandt PA, Goldbohm RA. A prospective cohort study on Allium vegetable consumption, garlic supplement use, and the risk of lung carcinoma in The Netherlands. *Cancer Res* 1994; **54**: 6148-6153 [PMID: 7954460]
51 **Dorgan JF**, Ziegler RG, Schoenberg JB, Hartge P, McAdams MJ, Falk RT, Wilcox HB, Shaw GL. Race and sex differences in associations of vegetables, fruits, and carotenoids with lung cancer risk in New Jersey (United States). *Cancer Causes Control* 1993; **4**: 273-281 [PMID: 8318643]
52 **McLaughlin JK**, Hrubec Z, Blot WJ, Fraumeni JF. Smoking and cancer mortality among U.S. veterans: a 26-year follow-up. *Int J Cancer* 1995; **60**: 190-193 [PMID: 7829214 DOI: 10.1002/ijc.2910600210]
53 **Enstrom JE**. Smoking cessation and mortality trends among two United States populations. *J Clin Epidemiol* 1999; **52**: 813-825 [PMID: 10529023]
54 **Kahn HA**. The Dorn study of smoking and mortality among U.S. veterans: report on eight and one-half years of observation. In: Haenszel W, editor. Epidemiological approaches to the study of cancer and other chronic diseases. Bethesda, MD: U.S. Department of Health, Education, and Welfare. Public Health Service National Cancer Institute, 1966: 1-125
55 **Engeland A**, Haldorsen T, Andersen A, Tretli S. The impact of smoking habits on lung cancer risk: 28 years' observation of 26,000 Norwegian men and women. *Cancer Causes Control* 1996; **7**: 366-376 [PMID: 8734831 DOI: 10.1007/BF00052943]
56 **Enstrom JE**, Heath CW. Smoking cessation and mortality trends among 118,000 Californians, 1960-1997. *Epidemiology* 1999; **10**: 500-512 [PMID: 10468422 DOI: 10.1097/00001648-199909000-00007]
57 **Gao CM**, Tajima K, Kuroishi T, Hirose K, Inoue M. Protective effects of raw vegetables and fruit against lung cancer among smokers and ex-smokers: a case-control study in the Tokai area of Japan. *Jpn J Cancer Res* 1993; **84**: 594-600 [PMID: 8340248 DOI: 10.1111/j.1349-7006.1993.tb02018.x]
58 **Gillis CR**, Hole DJ, Boyle P. Cigarette smoking and male lung cancer in an area of very high incidence. I. Report of a case-control study in the West of Scotland. *J Epidemiol Community Health* 1988; **42**: 38-43 [PMID: 3418284 DOI: 10.1136/jech.42.1.38]
59 **Haenszel W**, Shimkin MB, Mantel N. A retrospective study of lung cancer in women. *J Natl Cancer Inst* 1958; **21**: 825-842 [PMID: 13599015]
60 **Hammond EC**, Selikoff IJ, Seidman H. Asbestos exposure, cigarette smoking and death rates. *Ann N Y Acad Sci* 1979; **330**: 473-490 [PMID: 294198 DOI: 10.1111/j.1749-6632.1979.tb18749.x]
61 **Hammond EC**, Horn D. Smoking and death rates; report on forty-four monghs of follow-up of 187,783 men. II. Death rates by cause. *J Am Med Assoc* 1958; **166**: 1294-1308 [PMID: 13513355]
62 **Hirayama T**. Life-style and mortality: A large scale census based cohort study in Japan. In: Wahrendorf J, editor. Contributions to epidemiology and biostatistics. Basle: Karger, 1990: 6
63 **Hitosugi M**. Epidemiological study of lung cancer with special reference to the effect of air pollution and smoking habit. *Bull Inst Public Health* 1968; **17**: 237-256
64 **Gillis CR**, Hole DJ, Hawthorne VM. Cigarette smoking and

male lung cancer in an area of very high incidence. II. Report of a general population cohort study in the West of Scotland. *J Epidemiol Community Health* 1988; **42**: 44-48 [PMID: 3418285 DOI: 10.1136/jech.42.1.44]
65 **Humble CG**, Samet JM, Pathak DR, Skipper BJ. Cigarette smoking and lung cancer in 'Hispanic' whites and other whites in New Mexico. *Am J Public Health* 1985; **75**: 145-148 [PMID: 3966619 DOI: 10.2105/AJPH.75.2.145]
66 **Pathak DR**, Samet JM, Humble CG, Skipper BJ. Determinants of lung cancer risk in cigarette smokers in New Mexico. *J Natl Cancer Inst* 1986; **76**: 597-604 [PMID: 3457198]
67 **Friedman GD**, Tekawa I, Sadler M, Sidney S. Smoking and mortality: the Kaiser Permanente experience. In: Changes in cigarette-related disease risks and their implications for prevention and control. Rockville, MD: US Department of Health and Human Services, National Institutes of Health, National Cancer Institute, 1997: 477-499. Available from: URL: http://cancercontrol.cancer.gov/tcrb/monographs/8/m8_6.pdf
68 **Selby JV**, Friedman GD. Epidemiologic evidence of an association between body iron stores and risk of cancer. *Int J Cancer* 1988; **41**: 677-682 [PMID: 3366489 DOI: 10.1002/ijc.2910410507]
69 **Kanellakis A**, Trichopoulos D, Michalakopoulos N, Margoudakis S, Kanellaki K, Xirouchaki E, Kalapothaki V. The relationship between smoking of Greek cigarettes and the development of lung cancer. *Mater Med Greca* 1976; **4**: 351-355
70 **Trichopoulos D**, Kalandidi A, Tzonou A. Incidence and distribution of lung cancer in Greece. *Excerpta Med Int Congr Ser* 1982; **558**: 10-17
71 **Katsouyanni K**, Trichopoulos D, Kalandidi A, Tomos P, Riboli E. A case-control study of air pollution and tobacco smoking in lung cancer among women in Athens. *Prev Med* 1991; **20**: 271-278 [PMID: 2057473 DOI: 10.1016/0091-7435(91)90026-Z]
72 **Kaufman DW**, Palmer JR, Rosenberg L, Stolley P, Warshauer E, Shapiro S. Tar content of cigarettes in relation to lung cancer. *Am J Epidemiol* 1989; **129**: 703-711 [PMID: 2923118]
73 **Kinlen LJ**, Willows AN, Goldblatt P, Yudkin J. Tea consumption and cancer. *Br J Cancer* 1988; **58**: 397-401 [PMID: 3179194 DOI: 10.1038/bjc.1988.227]
74 **Knekt P**. Vitamin E and smoking and the risk of lung cancer. *Ann N Y Acad Sci* 1993; **686**: 280-27; discussion 280-27; [PMID: 8512253 DOI: 10.1111/j.1749-6632.1993.tb39187.x]
75 **Knekt P**, Seppänen R, Järvinen R, Virtamo J, Hyvönen L, Pukkala E, Teppo L. Dietary cholesterol, fatty acids, and the risk of lung cancer among men. *Nutr Cancer* 1991; **16**: 267-275 [PMID: 1775388 DOI: 10.1080/01635589109514165]
76 **Koo LC**, Ho JHC, Saw D. Is passive smoking an added risk factor for lung cancer in Chinese women? *J Exp Clin Cancer Res* 1984; **3**: 277-283
77 **Koo LC**, Ho JHC, Saw D. Active and passive smoking among female lung cancer patients and controls in Hong Kong. *J Exp Clin Cancer Res* 1983; **2**: 367-375
78 **Liaw KM**, Chen CJ. Mortality attributable to cigarette smoking in Taiwan: a 12-year follow-up study. *Tob Control* 1998; **7**: 141-148 [PMID: 9789932 DOI: 10.1136/tc.7.2.141]
79 **McDonald JC**, Liddell FD, Dufresne A, McDonald AD. The 1891-1920 birth cohort of Quebec chrysotile miners and millers: mortality 1976-88. *Br J Ind Med* 1993; **50**: 1073-1081 [PMID: 8280638 DOI: 10.1136/oem.50.12.1073]
80 **MacLennan R**, Da Costa J, Day NE, Law CH, Ng YK, Shanmugaratnam K. Risk factors for lung cancer in Singapore Chinese, a population with high female incidence rates. *Int J Cancer* 1977; **20**: 854-860 [PMID: 591126 DOI: 10.1002/ijc.2910200606]
81 **Matos E**, Vilensky M, Boffetta P, Kogevinas M. Lung cancer and smoking: a case-control study in Buenos Aires, Argentina. *Lung Cancer* 1998; **21**: 155-163 [PMID: 9857993 DOI: 10.1016/S0169-5002(98)00055-5]
82 **Lee PN**. Mortality from smoking-associated diseases in Great Britain. A statistical analysis of British data from the U.S.A.-U.K.-Norway migrant study. Sutton, Surrey: P N Lee Statistics and Computing Ltd., 1979. Available from: URL: http://www.pnlee.co.uk/Reports.htm
83 **Kuller LH**, Ockene JK, Meilahn E, Wentworth DN, Svendsen KH, Neaton JD. Cigarette smoking and mortality. MRFIT Research Group. *Prev Med* 1991; **20**: 638-654 [PMID: 1758843 DOI: 10.1016/0091-7435(91)90060-H]
84 **Kuller LH**, Ockene J, Meilahn E, Svendsen KH. Relation of forced expiratory volume in one second (FEV1) to lung cancer mortality in the Multiple Risk Factor Intervention Trial (MRFIT). *Am J Epidemiol* 1990; **132**: 265-274 [PMID: 2372006]
85 **Nam CB**, Hummer RA, Rogers RG. Underlying and multiple causes of death related to smoking. *Popul Res Policy Rev* 1994; **13**: 305-325 [DOI: 10.1007/BF01074340]
86 **Parkin DM**, Vizcaino AP, Skinner ME, Ndhlovu A. Cancer patterns and risk factors in the African population of southwestern Zimbabwe, 1963-1977. *Cancer Epidemiol Biomarkers Prev* 1994; **3**: 537-547 [PMID: 7827583]
87 **Pershagen G**, Akerblom G, Axelson O, Clavensjö B, Damber L, Desai G, Enflo A, Lagarde F, Mellander H, Svartengren M. Residential radon exposure and lung cancer in Sweden. *N Engl J Med* 1994; **330**: 159-164 [PMID: 8264737 DOI: 10.1056/NEJM199401203300302]
88 **Peto R**, Speizer FE, Cochrane AL, Moore F, Fletcher CM, Tinker CM, Higgins IT, Gray RG, Richards SM, Gilliland J, Norman-Smith B. The relevance in adults of air-flow obstruction, but not of mucus hypersecretion, to mortality from chronic lung disease. Results from 20 years of prospective observation. *Am Rev Respir Dis* 1983; **128**: 491-500 [PMID: 6614643]
89 **Pezzotto SM**, Poletto L. Occupation and histopathology of lung cancer: A case-control study in Rosario, Argentina. *Am J Ind Med* 1999; **36**: 437-443 [PMID: 10470008]
90 **Pezzotto SM**, Mahuad R, Bay ML, Morini JC, Poletto L. Variation in smoking-related lung cancer risk factors by cell type among men in Argentina: a case-control study. *Cancer Causes Control* 1993; **4**: 231-237 [PMID: 8391337]
91 **Prescott E**, Osler M, Hein HO, Borch-Johnsen K, Lange P, Schnohr P, Vestbo J. Gender and smoking-related risk of lung cancer. The Copenhagen Center for Prospective Population Studies. *Epidemiology* 1998; **9**: 79-83 [PMID: 9430273 DOI: 10.1097/00001648-199801000-00016]
92 **Segi M**, Kurihara M, Ishikawa S, Haenszel W. Epidemiological survey on lung cancer and smoking. *Lung Cancer* 1979; **19**: 157-165 [DOI: 10.2482/haigan.19.157]
93 **Shaw GL**, Falk RT, Deslauriers J, Frame JN, Nesbitt JC, Pass HI, Issaq HJ, Hoover RN, Tucker MA. Debrisoquine metabolism and lung cancer risk. *Cancer Epidemiol Biomarkers Prev* 1995; **4**: 41-48 [PMID: 7894323]
94 **Sobue T**, Suzuki T, Fujimoto I, Matsuda M, Doi O, Mori T, Furuse K, Fukuoka M, Yasumitsu T, Kuwahara O. Case-control study for lung cancer and cigarette smoking in Osaka, Japan: comparison with the results from Western Europe. *Jpn J Cancer Res* 1994; **85**: 464-473 [PMID: 8014103 DOI: 10.1111/j.1349-7006.1994.tb02381.x]
95 **Speizer FE**, Colditz GA, Hunter DJ, Rosner B, Hennekens C. Prospective study of smoking, antioxidant intake, and lung cancer in middle-aged women (USA). *Cancer Causes Control* 1999; **10**: 475-482 [PMID: 10530619 DOI: 10.1023/A:1008931526525]
96 **Stockwell HG**, Lyman GH, Waltz J, Peters JT. Lung cancer in Florida. Risks associated with residence in the central Florida phosphate mining region. *Am J Epidemiol* 1988; **128**: 78-84 [PMID: 2837899]
97 **Svensson C**, Pershagen G, Klominek J. Smoking and passive smoking in relation to lung cancer in women. *Acta Oncol* 1989; **28**: 623-629 [PMID: 2590538 DOI: 10.3109/02841868909092282]

98 **Tenkanen L**, Hakulinen T, Teppo L. The joint effect of smoking and respiratory symptoms on risk of lung cancer. *Int J Epidemiol* 1987; **16**: 509-515 [PMID: 3440661 DOI: 10.1093/ije/16.4.509]
99 **Tenkanen L**, Hakulinen T, Hakama M, Saxén E. Sauna, dust and migration as risk factors in lung cancer among smoking and non-smoking males in Finland. *Int J Cancer* 1985; **35**: 637-642 [PMID: 3997283]
100 **Tsugane S**, Watanabe S, Sugimura H, Arimoto H, Shimosato Y, Suemasu K. Smoking, occupation and family history in lung cancer patients under fifty years of age. *Jpn J Clin Oncol* 1987; **17**: 309-317 [PMID: 2826845]
101 **Tulinius H**, Sigfússon N, Sigvaldason H, Bjarnadóttir K, Tryggvadóttir L. Risk factors for malignant diseases: a cohort study on a population of 22,946 Icelanders. *Cancer Epidemiol Biomarkers Prev* 1997; **6**: 863-873 [PMID: 9367058]
102 **Tverdal A**, Thelle D, Stensvold I, Leren P, Bjartveit K. Mortality in relation to smoking history: 13 years' follow-up of 68,000 Norwegian men and women 35-49 years. *J Clin Epidemiol* 1993; **46**: 475-487 [PMID: 8501474 DOI: 10.1016/0895-4356(93)90025-V]
103 **Wakai K**, Ohno Y, Genka K, Ohmine K, Kawamura T, Tamakoshi A, Aoki R, Kojima M, Lin Y, Aoki K, Fukuma S. Smoking habits, local brand cigarettes and lung cancer risk in Okinawa, Japan. *J Epidemiol* 1997; **7**: 99-105 [PMID: 9255031 DOI: 10.2188/jea.7.99]
104 **Wu AH**, Henderson BE, Pike MC, Yu MC. Smoking and other risk factors for lung cancer in women. *J Natl Cancer Inst* 1985; **74**: 747-751 [PMID: 3857370]
105 **Harris RE**, Zang EA, Anderson JI, Wynder EL. Race and sex differences in lung cancer risk associated with cigarette smoking. *Int J Epidemiol* 1993; **22**: 592-599 [PMID: 8225730 DOI: 10.1093/ije/22.4.592]
106 **Yamaguchi N**, Kido M, Hoshuyama T, Manabe H, Kikuchi Y, Nishio T, Ohshima LH, Watanabe S. A case-control study on occupational lung cancer risks in an industrialized city of Japan. *Jpn J Cancer Res* 1992; **83**: 134-140 [PMID: 1555994 DOI: 10.1111/j.1349-7006.1992.tb00077.x]

**P- Reviewers** Iftikhar IH, Shen XC, Zhang J **S- Editor** Gou SX
**L- Editor** A **E- Editor** Zheng XM

World Journal of Meta-Analysis

Online Submissions: http://www.wjgnet.com/esps/
wjma@wjgnet.com
doi:10.13105/wjma.v1.i2.78

*World J Meta-Anal* 2013 August 26; 1(2): 78-82
ISSN 2308-3840 (online)


META-ANALYSIS

# Ophthalmic adverse drug reactions: A nationwide detection using hospital databases

Ana Miguel, Filipe Henriques, Bernardo Marques, Joana Marques, Alberto Freitas, Fernando Lopes, Luís Azevedo, Altamiro Costa Pereira

Ana Miguel, Bernardo Marques, Alberto Freitas, Fernando Lopes, Luís Azevedo, Altamiro Costa Pereira, Centre for Research in Health Technologies and Information Systems, Faculty of Medicine, University of Porto, 4430-182 V.N.Gaia, Portugal
Ana Miguel, Filipe Henriques, Department of Ophthalmology, Central University Hospital of Coimbra, 3040-091 Coimbra, Portugal
Joana Marques, Northern Pharmacovigilance Centre, Faculty of Medicine, University of Porto, 4430-182 V.N.Gaia, Portugal
Author contributions: All authors contributed to this paper.
Correspondence to: Ana Miguel, MD, Centre for Research in Health Technologies and Information Systems, Faculty of Medicine, University of Porto, Rua Quinta do Sardoal, VE3, nº10, 4430-182 V.N.Gaia, Portugal. myworld_ana@hotmail.com
Telephone: +35-193-2482477 Fax: +35-123-9119864
Received: February 4, 2013 Revised: June 10, 2013
Accepted: July 18, 2013
Published online: August 26, 2013

## Abstract

**AIM:** To detect ophthalmic adverse drug reactions (ADRs), that occurred in Portugal from 2000 to 2009, through the utilization of administrative hospital databases. We also intended to compare the results of this methodology with spontaneous reporting.

**METHODS:** We conducted a retrospective nationwide study using hospital administrative databases, which included all inpatients and outpatients in all public hospitals in Portugal, from 2000 to 2009. We used International Classification of Diseases - 9th Revision - Clinical Modification (ICD-9-CM) coding data that allowed the detection of ADRs. We used WHO's definition for ADR. We searched all of ICD-9-CM terms in Ophthalmology for codes that included "drug-induced", "iatrogenic", "toxic" and all other that could signal an ADR, such as "362.55 - toxic maculopathy" or "365.03 - steroid responders", and also "E" codes (codes from E930 to E949.9, that exclude intoxications and errors).

**RESULTS:** From 11944725 hospitalizations or ambulatory episodes within that period of time, we identified 1524 probable ophthalmic ADRs (corresponding to a frequency of 1.28 per 10000 episodes) and an additional 100 possible ophthalmic ADRs. We used only 4 person-hours in the application of this methodology. A total of 113 spontaneous reports arose from ophthalmic ADRs from 2000 to 2009 in Portugal (frequency of 0.095 per 10000 episodes).To our knowledge, this was the first estimate of the frequency of ophthalmic ADRs through the use of databases, and the first nationwide estimate of ophthalmic ADRs, in Portugal. We identified 1524 probable ADRs and 100 possible ADRs.

**CONCLUSION:** This database methodology adapted for Ophthalmology may represent a new approach for the detection of ophthalmic ADRs, since these codes exist in the ICD-9-CM classification. Its performance was clearly superior to spontaneous reporting.



**Key words:** Adverse drug reactions; Ophthalmology; Ocular; Databases; Pharmacovigilance

**Core tip:** We used International Classification of Diseases - 9th Revision - Clinical Modification coding data for the detection of adverse drug reactions (ADRs). From 11944725 episodes, we identified 1524 probable ophthalmic ADRs. 113 spontaneous reports arose from that population. This was the first nationwide study of ophthalmic ADRs and may represent a new Pharmacovigilance approach, with a higher detection than spontaneous reporting.

Miguel A, Henriques F, Marques B, Marques J, Freitas A, Lopes F, Azevedo L, Pereira AC. Ophthalmic adverse drug reactions: A nationwide detection using hospital databases. *World J Meta-Anal* 2013; 1(2): 78-82 Available from: URL: http://www.

wjgnet.com/2308-3840/full/v1/i2/78.htm DOI: http://dx.doi.org/10.13105/wjma.v1.i2.78

## INTRODUCTION

Adverse drug reactions (ADRs) are responsible for significant morbidity, mortality and costs in Health Care systems[1]. They may occur in 16.9% of patients during hospitalization (95%CI: 13.5-20.2)[2] and provoke 5.3% of hospital admissions (interquartile range 2.7%-9.0%)[3]. ADRs are a frequent cause of death in developed countries[4]. However, in Ophthalmology the evidence is scarce and lacks systematization[5]. A review about challenges in ADRs in Ophthalmology[5] concluded that there are several areas that can be improved, namely by applying always the definition of ADR of the World Health Organization (WHO)[6], by performing a causality assessment in each ADR (which determines the probability of representing a true ADR; the most utilized causality assessments of ADRs are from WHO[7] and from Naranjo *et al*[8]).

The development and validation of new methodologies for an improved detection of ADRs would be another area of improvement[5,9]. There are Pharmacovigilance methodologies[9] used for the detection of ADRs and that can be adapted for detecting ADRs in Ophthalmology, but they may have methodological issues: Spontaneous reporting is the most used (it needs low resources) and is the only Pharmacovigilance method continuously used in the majority of countries, being the main support of WHO International Drug Program. However, it has several limitations, namely, the smallest detection rate of several Pharmacovigilance methods[10], under-reporting[11], heterogeneous report quality[12] and increased risk of bias[12]. Intensive and prospective monitoring are methodologies with good detection rates but too resource-consuming for continuous application[13].

Administrative hospital databases have large clinical information and thus may represent an interesting Pharmacovigilance approach with readily available and cheap information[10]. Some authors have utilized databases[10,14] for the detection of ADRs, taking advantage of the large quantity of clinical information readily available, containing coding data that can be used as an alert for the detection of an ADR, with low relatively low resources required.

Our purpose was to identify and characterize ophthalmic ADRs in a Nationwide study in Portugal, using hospital databases with clinical information.

## MATERIALS AND METHODS

### Study design

A retrospective study was performed for ADR identification using hospital administrative databases with information from all public hospitals in Portugal, from 2000 to 2009, obtained from our National Health Department (data from the second semester of 2009 was not available). These databases contain anonymized data for patient identification, episode and process number, and also information on age, sex, admission date, discharge date, ward(s), hospital attended (tertiary, university), area of Healthcare, district, outcome (death, discharge, transfer), payment data and International Classification of Diseases - 9th Revision - Clinical Modification (ICD-9-CM)[15] codes for: diagnoses (principal diagnosis, other diagnosis up to 19), procedures (up to 20) and external causes (up to 20). Patient population included all patients hospitalized or admitted for ambulatory care, in all public hospitals in Portugal, from 2000 to 2009 (inpatients and outpatients). All investigations were performed according to the guidelines of the Declaration of Helsinki and Institutional Review Board approval from was obtained.

### Definition of ADR

There is some misuse of terms in this matter; therefore we present definitions.

An ADR[6] is: "any noxious, unintended and undesired effect of a drug, which occurs at doses used in humans for prophylaxis, diagnosis, or therapy". Therefore, to increase specificity, we wanted to assess only ADRs. Adverse drug event is not a synonym of ADR. There are other definitions of ADR, namely from Karch *et al*[16] and from Edwards *et al*[17], but we used the definition of WHO. An adverse event[18] is: "an injury related to medical management (all aspects of care, including diagnosis and treatment, failure to diagnose or treat, and the systems and equipment used to deliver care), in contrast to complications of disease". An adverse drug event[19] is: "An injury related to the use of a drug, although the causality of this relationship may not be proven". These include medication errors (namely the prescription of a wrong dose) and ADRs. We aimed to assess strictly ADRs.

### Detection of ADRs

Hospital administrative databases include information of diagnosis. Codes searched for ADR identification were adapted to the specificities of Ophthalmology and resulted from a thorough search of: all terms of ICD-9-CM in Ophthalmology that included "drug-induced", "iatrogenic", "toxic" and all codes that could signal an ADR, such as "362.55 - toxic maculopathy" or "365.03 - steroid responsers", as detailed in the Results Section.

We also performed a search of general ADRs through the use of 'E' codes (ICD-9-CM codes from E930 to E949.9, designed to represent ADRs and already excluding wrong doses, errors and intoxications) to assess if these general ADRs could detect ophthalmic ADRs.

In this study, we performed a query of Ophthalmology in a nationwide study using administrative databases, including inpatients and ambulatory patients. Our main outcome was ADR detection. Secondary outcomes included: type of ADR, age, sex, admission diagnosis, other diagnoses, hospital length-of-stay and year of discharge. We performed WHO's causality assessments of ADRs,

with two independent reviewers. Differences were resolved by consensus. A third review was consulted to help resolved differences. We also registered how many person-hours were spent in the application of this methodology, to estimate cost (resources spent). The number of person-hours refers to the number of hours and number of people used in the application of this methodology; commonly used in the comparison of different Pharmacovigilance methodologies[19]. The number of spontaneous reporting of ADRs in hospitalized patients from 2000 to 2009 was obtained from Portuguese National Authority of Medicines (INFARMED), for comparison[20].

### Statistical analysis

Statistical analyses were done using the $\chi^2$ test for categorical variables (or exact Fisher's test whenever possible), Student's *t*-test for normally distributed continuous variables and Mann-Whitney or Kruskal-Wallis for variables without normal distribution, using SPSS v20. The a priori level of significance was $P < 0.05$.

## RESULTS

### Study population

There were 11944725 patients hospitalized or with ambulatory episodes in public hospitals of Portugal, from 2000 to the first semester of 2009. The baseline characteristics of the study population ($n$ = 11944725) are shown in Table 1. The mean age of hospitalized patients was 48 ± 27 years and in 55.2% of episodes the patient was female. We spent only 4 person-hours in the application of this methodology.

From 2000, there was a slight increase in the number of hospitalizations in Portugal. Specific ophthalmic ADRs ($n$ = 1524) were detected through the search of codes that could represent particular ophthalmic ADRs, as shown in Table 2. This corresponds to a frequency of 1.28 ophthalmic ADR per 10000 episodes. Additionally, 100 episodes that could possibly correspond to an ophthalmic ADR were also detected (Table 2). Therefore, a total of 1624 possible ophthalmic ADRs were detected. These possible ADRs included: conjunctival concretions, pigmentations and deposits (which can be caused by drugs such as topical adrenaline[21], but also by other factors, therefore may correspond to an ADR in some cases) and acquired color vision deficiencies (which may be caused by drugs such as sildenafil[22], but have other non related causes).

The search of general ADRs through the use of "E" codes allowed us to identify 116720 ADRs, but only 62 of them corresponded to the ophthalmic ADRs that were identified.

The total number of spontaneous notifications of ADRs in Portugal from 2000 to 2009 was 13562, from which 113 were spontaneous reports specific of ophthalmic ADRs. There were 553 additional spontaneous reports of systemic ADRs that included some ophthalmic manifestations.

**Table 1 Socio-demographic characteristics of study population**

| Characteristic | Value |
|---|---|
| Number of episodes (inpatient, ambulatory) | 11944725 |
| Mean age (yr, mean ± SD) | 48 ± 27 |
| Female gender n (%) | 6598266 (55.2) |
| District with higher number of hospitalizations | 1st: Lisbon 21.2%<br>2nd: Oporto 17.2%<br>3rd: Setubal 7.66% |
| Mean hospital length-of-stay for inpatients (d, mean ± SD) | 7.1 ± 3.21 |
| Number of probable ophthalmic ADRs | 1524 |

ADRs: Adverse drug reactions.

**Table 2 Clinical codes searched and respective results in the portuguese database**

| ICD-9-CM code | Diagnosis | No. of episodes |
|---|---|---|
| Specific ophthalmic ADR codes | | |
| 362.55 | Toxic maculopathy | 1388 |
| 365.03 | Steroid responders | 4 |
| 365.31, 365.32 | Corticosteroid-induced glaucoma | 0 |
| 364.55 | Miotic pupillary cyst (provoked by pilocarpine) | 2 |
| 364.81 | Floppy iris syndrome | 2 |
| 366.45 | Toxic cataract | 83 |
| 367.89 | Other drug-induced disorders of refraction and accommodation, Toxic disorders of refraction and accommodation | 25 |
| 377.34 | Toxic optic neuropathy, Toxic amblyopia | 20 |
| Possible signs of ophthalmic ADRs | | |
| 366.46 | Cataract associated with radiation and other physical influences | 10 |
| 372.54 | Conjunctival concretions | 67 |
| 372.55 | Conjunctival pigmentations, including conjunctival argyrosis | |
| 372.56 | Conjunctival deposits | |
| 368.55 | Acquired color vision deficiencies | 23 |
| 368.59 | Other color vision deficiencies | |
| | Sub-Total specific | 1524 |
| | Total | 1624 |

ICD-9-CM: Classification of Diseases - 9th Revision - Clinical Modification; ADRs: Adverse drug reactions.

## DISCUSSION

To our knowledge, this is the first estimate of the frequency of ophthalmic ADRs through the use of administrative databases, and the first to apply a nationwide estimate of ophthalmic ADRs, in Portugal. We identified 1524 probable ADRs and 100 possible ADRs. This may represent a new approach for the detection of ophthalmic ADRs, since these codes exist in the ICD-9-CM classification.

The strengths of our study include: our comprehensive database, which contains data from all hospitalizations and ambulatory episodes in every public hospital in Portugal within almost a decade, the fact that this is a new methodology to aid ADR detection (until now only case reports and spontaneous reports were available for

ADR detection), and the fact that these codes are widely available and universal, making possible to easily build estimates of ophthalmic ADRs in other countries and other years. In fact, it would be very interesting to see if ophthalmic ADRs in Portugal have the same distribution, frequency and characteristics in comparison with other countries, therefore further studies are necessary.

Limitations of our work are inherent to the use of administrative databases, which may contain incomplete or wrong data and coding bias[23] (in which coders select a different code to increase reimbursement to their hospital). The small number of ADRs found may be considered a limitation, but on the other hand this is a methodology resource-sparing (only 4 person-hours spent in its application), having potential for widespread application in other countries. Also, this method identified 1524 probable ADRs, a much higher number than the number of ophthalmic ADRs found by spontaneous reporting: 113.

We suggest complementing spontaneous reporting with this database methodology to increase detection of ophthalmic ADRs. In fact, the complementary use of several methodologies is defended by several authors[24], in order to enhance ADR detection and increase patient safety. Finally, we believe that after this study, these codes should be applied prospectively in a future study in a nation-wide basis, enabling an expert to confirm each ADR and causing drug, to further complete and validate the data suggested here, and to integrate this method as a Pharmacovigilance methodology.

In conclusion, Ophthalmology represents simultaneously a challenge and an opportunity to identify ADRs. This is the first nationwide estimate of ophthalmic ADRs. Administrative databases are a useful methodology for the detection of ocular ADRs, but require adapted diagnoses codes. They may underestimate the real number of ADRs, but nevertheless they have the potential to complement spontaneous reporting as a methodology for ophthalmic ADR detection, with a higher detection rate.

## ACKNOWLEDGMENTS

The authors would like to thank ACSS for providing access to the data, and express gratitude to the statistical support given by the research project HR-QoD - Quality of data (outliers, inconsistencies and errors) in hospital inpatient databases: methods and implications for data modeling, cleansing and analysis (project PTDC/SAU-ESA/75660/2006). The authors would also like to thank the INFARMED, Portuguese National Authority of Medicines and Health Products, for the data kindly provided about spontaneous reporting in Portugal.

## COMMENTS

### Background

Adverse drug reactions (ADRs) are a frequent cause of death in developed countries. However, in Ophthalmology the evidence is scarce and lacks systematization.

### Research frontiers

There are Pharmacovigilance methodologies used for the detection of ADRs and that can be adapted for detecting ADRs in Ophthalmology, but they may have methodological issues.

### Innovations and breakthroughs

This is the first estimate of the frequency of ophthalmic ADRs through the use of administrative databases, and the first to apply a nationwide estimate of ophthalmic ADRs, in Portugal.

### Applications

The authors suggest complementing spontaneous reporting with this database methodology to increase detection of ophthalmic ADRs.

### Peer review

This is a well written article reporting the adverse effects of ophthalmic drugs. The methods are well described, and the results are easy to understand.

## REFERENCES

1 **Davies EC**, Green CF, Mottram DR, Pirmohamed M. Adverse drug reactions in hospital in-patients: a pilot study. *J Clin Pharm Ther* 2006; **31**: 335-341 [PMID: 16882102]
2 **Miguel A**, Azevedo LF, Araújo M, Pereira AC. Frequency of adverse drug reactions in hospitalized patients: a systematic review and meta-analysis. *Pharmacoepidemiol Drug Saf* 2012; **21**: 1139-1154 [PMID: 22761169 DOI: 10.1002/pds.3309]
3 **Kongkaew C**, Noyce PR, Ashcroft DM. Hospital admissions associated with adverse drug reactions: a systematic review of prospective observational studies. *Ann Pharmacother* 2008; **42**: 1017-1025 [PMID: 18594048]
4 **Lazarou J**, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *JAMA* 1998; **279**: 1200-1205 [PMID: 9555760]
5 **Fraunfelder FW**, Fraunfelder FT. Scientific challenges in postmarketing surveillance of ocular adverse drug reactions. *Am J Ophthalmol* 2007; **143**: 145-149 [PMID: 17188050]
6 **World Health Organization**. Adverse Drug Reaction Monitoring. Available from: URL: http://www.who.int/medicines/areas/quality_safety/safety_efficacy/advdrugreactions/en/index.html
7 **World Health Organization**. The Importance of Pharmacovigilance - Safety Monitoring of Medicinal Products. Available from: URL: http://apps.who.int/medicinedocs/en/d/Js4893e/
8 **Naranjo CA**, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, Janecek E, Domecq C, Greenblatt DJ. A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther* 1981; **30**: 239-245 [PMID: 7249508]
9 **Davies EC**, Green CF, Mottram DR, Pirmohamed M. Adverse drug reactions in hospitals: a narrative review. *Curr Drug Saf* 2007; **2**: 79-87 [PMID: 18690953]
10 **Miguel A**, Azevedo LF, Lopes F, Freitas A, Pereira AC. Methodologies for the detection of adverse drug reactions: comparison of hospital databases, chart review and spontaneous reporting. *Pharmacoepidemiol Drug Saf* 2013; **22**: 98-102 [PMID: 23027707 DOI: 10.1002/pds.3348]
11 **Figueiras A**, Herdeiro MT, Polónia J, Gestal-Otero JJ. An educational intervention to improve physician reporting of adverse drug reactions: a cluster-randomized controlled trial. *JAMA* 2006; **296**: 1086-1093 [PMID: 16954488 DOI: 10.1001/jama.296.9.1086]
12 **Bandekar MS**, Anwikar SR, Kshirsagar NA. Quality check of spontaneous adverse drug reaction reporting forms of different countries. *Pharmacoepidemiol Drug Saf* 2010; **19**: 1181-1185 [PMID: 20845409]
13 **Pourseyed S**, Fattahi F, Pourpak Z, Gholami K, Shariatpanahi SS, Moin A, Kazemnejad A, Moin M. Adverse drug reactions in patients in an Iranian department of internal medicine. *Pharmacoepidemiol Drug Saf* 2009; **18**: 104-110 [PMID: 19101919]
14 **van der Hooft CS**, Sturkenboom MC, van Grootheest K,

Kingma HJ, Stricker BH. Adverse drug reaction-related hospitalisations: a nationwide study in The Netherlands. *Drug Saf* 2006; **29**: 161-168 [PMID: 16454543]
15 **Slee VN**. The International Classification of Diseases: ninth revision (ICD-9). *Ann Intern Med* 1978; **88**: 424-426 [PMID: 629506]
16 **Karch FE**, Lasagna L. Adverse drug reactions. A critical review. *JAMA* 1975; **234**: 1236-1241 [PMID: 1242749]
17 **Edwards IR**, Aronson JK. Adverse drug reactions: definitions, diagnosis, and management. *Lancet* 2000; **356**: 1255-1259 [PMID: 11072960]
18 **World Health Organization**. WHO Draft Guidelines for Adverse Event Reporting and Learning Systems. 2005. Available from: URL: http://www.who.int/patientsafety/events/05/Reporting_Guidelines.pdf
19 **Nebeker JR**, Barach P, Samore MH. Clarifying adverse drug events: a clinician's guide to terminology, documentation, and reporting. *Ann Intern Med* 2004; **140**: 795-801 [PMID: 15148066]
20 INFARMED - National Portuguese Authority of Drug and Health Products. Available from: URL: http://www.infarmed.pt
21 **Fong DS**, Frederick AR, Richter CU, Jakobiec FA. Adrenochrome deposit. *Arch Ophthalmol* 1993; **111**: 1142-1143 [PMID: 8352696]
22 **Azzouni F**, Abu samra K. Are phosphodiesterase type 5 inhibitors associated with vision-threatening adverse events? A critical analysis and review of the literature. *J Sex Med* 2011; **8**: 2894-2903 [PMID: 21771280]
23 **Seiber EE**. Physician code creep: evidence in Medicaid and State Employee Health Insurance billing. *Health Care Financ Rev* 2007; **28**: 83-93 [PMID: 17722753]
24 **Whitstock MT**, Pearce CM, Ridout SC, Eckermann EJ. Using clinical trial data and linked administrative health data to reduce the risk of adverse events associated with the uptake of newly released drugs by older Australians: a model process. *BMC Public Health* 2011; **11**: 361 [PMID: 21600026]

**P- Reviewers** Onakpoya I, Saokaew S **S- Editor** Zhai HH
**L- Editor** A **E- Editor** Zheng XM

World Journal of Meta-Analysis

Online Submissions: http://www.wjgnet.com/esps/
wjma@wjgnet.com
doi:10.13105/wjma.v1.i2.83

*World J Meta-Anal* 2013 August 26; 1(2): 83-89
ISSN 2308-3840 (online)




# Prevalence of hypertension in India: A meta-analysis

Tanu Midha, Bhola Nath, Ranjeeta Kumari, Yashwant Kumar Rao, Umeshwar Pandey

Tanu Midha, Department of Community Medicine, GSVM Medical College, Kanpur 208002, India
Bhola Nath, Department of Community Medicine, VCSG Government Medical College, Uttarakhand 246174, India
Ranjeeta Kumari, Department of Community Medicine, AIIMS, Rishikesh 249201, India
Yashwant Kumar Rao, Department of Pediatrics, GSVM Medical College, Kanpur 208002, India
Umeshwar Pandey, Department of Cardiology, LPS Institute of Cardiology and Cardiac Surgery, Kanpur 208002, India
Author contributions: Midha T, Rao YK, Kumari R and Nath B designed the review; Midha T, Rao YK, Kumari R, Nath B and Pandey U collected the data; Midha T, Rao YK, Kumari R and Nath B analyzed the data; Midha T, Rao YK, Kumari R, Nath B and Pandey U wrote the paper.
Correspondence to: Tanu Midha, Assistant Professor, Department of Community Medicine, GSVM Medical College, Swaroop Nagar, Kanpur 208002, India. tanumidha2001@gmail.com
Telephone: +91-512-2216721 Fax: +91-512-2535483
Received: April 24, 2013 Revised: July 18, 2013
Accepted: August 4, 2013
Published online: August 26, 2013

## Abstract

**AIM:** To determine the prevalence of hypertension in the urban and rural population of India.

**METHODS:** Relevant studies were identified through computer based and manual searches using MEDLINE/PubMed, Google scholar, EMBASE, Cochrane Library and reference lists of prevalence studies from January 2000 to June 2012. A total of 12 studies were included in the meta-analysis of hypertension in urban India and 10 studies in the analysis of hypertension in rural India after applying the inclusion and exclusion criteria. Estimates of prevalence were calculated using the random effect model for meta-analysis.

**RESULTS:** The electronic search using appropriate keywords identified 177 titles for prevalence of hypertension in urban India, of which 165 were excluded, and 133 titles for prevalence in rural India, of which 123 were excluded after applying the inclusion criteria. Twelve studies including 125333 subjects were analyzed to assess the prevalence of hypertension in the urban Indian population, whereas ten studies including 24800 subjects were analyzed to determine the prevalence of hypertension in the rural Indian population. The prevalence of hypertension in the urban population was estimated to be 40.8% (95%CI: 40.5%-41.0%) and that of hypertension in the rural population was 17.9% (95%CI: 17.5%-18.3%). It is evident that the prevalence of hypertension is significantly higher in the urban population of India compared to the rural.

**CONCLUSION:** Current evidence suggests that policies and interventions should be prioritized for reduction of hypertension in the adult Indian population, especially the urban population.



**Key words:** Prevalence; Hypertension; Meta-analysis; India; Urban; Rural

**Core tip:** A meta-analysis of prevalence studies on hypertension in India from January 2000 to June 2012 reveals a high prevalence of hypertension in the urban (40.8%) as well as rural population (17.9%). The prevalence of hypertension is markedly higher in the urban population compared to the rural population, but the prevalence in the rural population is also a matter of concern with almost every fifth individual at risk. This is indicative of the epidemiological transition, which must raise an alarm for policy makers and health care professionals. Primordial and primary prevention of hypertension can bring about a substantial reduction in cardiovascular morbidity and mortality which occurs as a consequence of hypertension.

Midha T, Nath B, Kumari R, Rao YK, Pandey U. Prevalence of hypertension in India: A meta-analysis. *World J Meta-Anal* 2013; 1(2): 83-89 Available from: URL: http://www.wjgnet.

com/2308-3840/full/v1/i2/83.htm DOI: http://dx.doi.org/10.13105/wjma.v1.i2.83

## INTRODUCTION

Globally, the overall prevalence of hypertension or raised blood pressure in adults aged 25 and above was around 40% in 2008[1]. Worldwide, hypertension is estimated to cause 7.5 million deaths, about 12.8% of the total deaths. Hypertension accounts for 57 million disability adjusted life years (DALYS) or 3.7% of total DALYS[1]. The World Health Organization (WHO) has estimated that globally about 62% of cerebrovascular diseases and 49% of ischemic heart diseases are attributable to suboptimal blood pressure (systolic > 115 mmHg), with little variation by sex[2]. One in three adults worldwide has high blood pressure. Hypertension increases the risk of heart attack, stroke, kidney failure and many other associated co morbidities. Treating raised blood pressure and maintaining it below 140/90 mmHg is associated with a reduction in cardiovascular complications[1].

The theme for World Health Day (WHD) 2013 is "high blood pressure"[3]. The goal of WHD 2013 is to reduce heart attacks and strokes. Keeping in line with the WHO-Government of India Country Cooperation Strategy, the WHD 2013 events in India are aimed at raising the awareness amongst national policymakers, program managers and other stakeholders on the need to strengthen the Indian health system to make it competent enough to respond to hypertension and related co morbidities[3].

Hypertension is a controllable disease and it has been reported that targeted reductions in people with hypertension are expected to produce large reductions in the burden of cardiovascular disease[4]. According to the seventh report of the Joint National Committee (JNC-7) on prevention, detection, evaluation and treatment of high blood pressure, adoption of healthy lifestyles by all individuals is critical for the prevention of high blood pressure[5]. Accurate estimates of hypertension are therefore necessary to plan effective control measures.

A meta-analysis showed an increase in the prevalence of hypertension in India over the years from 1%-3% in 1950 to 10%-30.9% in 2002[6]. Another cause for concern is the epidemiological transition, as it is likely that the prevalence of risk factors, and consequently the prevalence of hypertension and cardiovascular diseases, would rise with the socioeconomic development of rural areas in India.

India accounts for 17% of the world's population, the second largest in the world, and hence it contributes largely to the statistics of any disease in the world[7]. Given the fact that hypertension is on the rise in developing countries like India, this meta-analysis was designed to consolidate the available data to find out the current prevalence of hypertension in urban and rural India.

## MATERIALS AND METHODS

### Search strategy

We searched MEDLINE/PubMed, Google scholar, EMBASE, Cochrane Library and reference lists of prevalence studies from January 2000 to June 2012. Internet searches used permutations of medical subject headings for prevalence studies on hypertension in India. The following keywords were looked for individually or in association: hypertension, India, prevalence, blood pressure, systolic, diastolic, mmHg. The limits included were: English for the language category and humans for the study category (Figure 1).

### Selection criteria

The studies that met all of the following criteria were included in the present meta-analysis: (1) they were prevalence studies; (2) the study design was cross-sectional; (3) the age group included in the study was 20 years and above; (4) the study was conducted in the Indian population; (5) the cut-off for classification of hypertension was systolic blood pressure $\geqslant$ 140 mmHg and/or diastolic blood pressure $\geqslant$ 90 mmHg; and (6) the study contained original data. All the studies had a cross-sectional design and blood pressure measurement on a single visit was considered. Exclusion criteria: reviews, letters to editors, case series and case-control studies were not included because of insufficient data for analysis.

### Statistical analysis

Data was analyzed using the statistical software Comprehensive Meta-analysis V2. The random effect model was used to calculate the estimate of the prevalence of hypertension rather than the fixed effect model. The random effect model takes into account any heterogeneity inherent in the meta-analysis.

## RESULTS

### Literature review

The electronic search in Pubmed, using the keywords "prevalence of hypertension in urban India", identified 177 titles, of which 165 were excluded based on review of titles, abstracts and text after applying the inclusion criteria. To avoid bias due to selection criteria and blood pressure criteria used in various studies, age, blood pressure cut-off and study design criteria were taken into consideration. The remaining 12 studies were included in the analysis. Similarly, following electronic search in Pubmed, using the keywords "prevalence of hypertension in rural India", we identified 133 titles, of which 123 were excluded based on review of titles, abstracts and text after applying the inclusion criteria. The remaining 10 were included in the analysis.

### Study characteristics

In total, 12 studies were included in the meta-analysis of hypertension in urban India and 10 studies in the analysis of hypertension in rural India.

**Table 1 Review of studies on the prevalence of hypertension**

| First author | Place | Age group (yr) | Sample size | Prevalence |
|---|---|---|---|---|
| In the urban Indian population | | | | |
| Anand[10] | Maharashtra | 30-60 | 1662 | 34.0 |
| Gupta *et al*[11] | Rajasthan | > 20 | 1123 | 33.4 |
| Shanthirani *et al*[12] | Tamil Nadu | > 20 | 1262 | 21.1 |
| Gupta *et al*[18] | Maharashtra | > 35 | 88653 | 47.9 |
| Prabhakaran *et al*[16] | Delhi | 20-59 | 2935 | 30.0 |
| Reddy *et al*[13] | Multi-centric | 20-69 | 19973 | 27.7 |
| Mohan *et al*[14] | Tamil Nadu | > 20 | 2350 | 20.0 |
| Kaur *et al*[15] | Tamil Nadu | 18-69 | 2262 | 27.2 |
| Yadav *et al*[33] | Uttar Pradesh | > 30 | 1746 | 32.2 |
| Midha *et al*[17] | Uttar Pradesh | > 20 | 400 | 32.8 |
| Gupta[27] | Multi-centric | 35-70 | 926 | 48.2 |
| Chakraborty *et al*[9] | West Bengal | 18-60 | 433 | 17.6 |
| In the rural Indian population | | | | |
| Kusuma *et al*[31] | Andhra Pradesh | > 20 | 1316 | 21.0 |
| Hazarika *et al*[30] | Assam | > 30 | 3180 | 33.3 |
| Midha *et al*[17] | Uttar Pradesh | > 20 | 400 | 14.5 |
| Todkar *et al*[19] | Maharashtra | > 20 | 1297 | 7.2 |
| Bhardwaj *et al*[24] | Himachal Pradesh | > 18 | 1092 | 35.9 |
| Kinra *et al*[32] | Multi-centric | 20-69 | 1983 | 20.0 |
| Rajasekar *et al*[20] | Tamil Nadu | > 30 | 1905 | 19.1 |
| Kadu *et al*[21] | Maharashtra | > 18 | 2196 | 12.8 |
| Bansal *et al*[22] | Uttarakhand | > 15 | 968 | 32.3 |
| Kaur *et al*[23] | Tamil Nadu | 25-64 | 10463 | 21.4 |

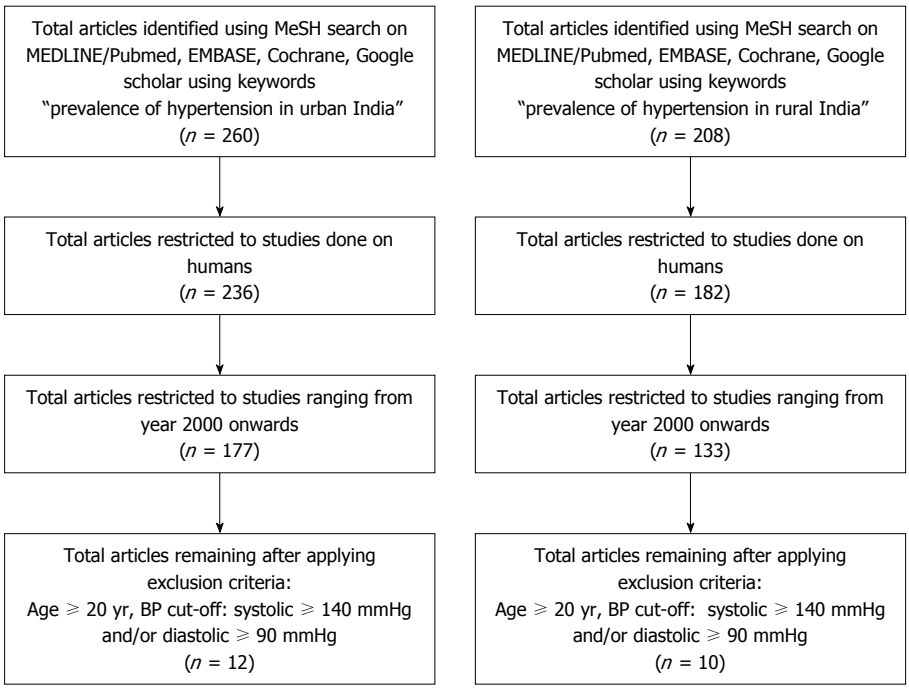



**Figure 1 Flow diagram of selection process.** Course of systematic literature review on prevalence of hypertension in urban and rural India. BP: Blood pressure.

### *Meta-analysis*

After analysis of 125333 subjects from twelve studies, the prevalence of hypertension in the urban Indian population was found to be 40.8% (95%CI: 40.5%-41.0%) (Table 1). Ten studies including 24800 subjects were analysed and the prevalence of hypertension in the rural population was estimated to be 17.9% (95%CI: 17.5%-18.3%) (Table 1). Figure 2 shows the meta-analysis of prevalence of hypertension in the urban and rural Indian population respectively. The overall prevalence rate is represented by the random effect size which was estimated to be 40.8% in urban and 17.9% in the rural population.

## DISCUSSION

In the present meta-analysis, the prevalence of hypertension was estimated to be 40.8% in urban and 17.9% in

**A**

Meta Analysis

| Study name | Statistics for each study | | | | | | |
|---|---|---|---|---|---|---|---|
| | Point estimate | SE | Variance | Lower limit | Upper limit | *Z*-value | *P*-value |
| Anand[10] | 34.00 | 1.160 | 1.34 | 31.72 | 36.27 | 29.310 | 0.000 |
| Gupta *et al*[11] | 33.40 | 1.410 | 1.98 | 30.63 | 36.16 | 23.688 | 0.000 |
| Shanthirani *et al*[12] | 21.10 | 1.150 | 1.32 | 18.84 | 23.35 | 18.348 | 0.000 |
| Gupta *et al*[18] | 47.90 | 0.170 | 0.02 | 47.56 | 48.23 | 281.765 | 0.000 |
| Prabhakaran *et al*[16] | 30.00 | 0.850 | 0.72 | 28.33 | 31.66 | 35.294 | 0.000 |
| Reddy *et al*[13] | 27.20 | 0.310 | 0.09 | 26.59 | 27.80 | 87.742 | 0.000 |
| Mohan *et al*[14] | 20.00 | 0.830 | 0.68 | 18.37 | 21.62 | 24.096 | 0.000 |
| Kaur *et al*[15] | 27.20 | 0.940 | 0.88 | 25.35 | 29.04 | 28.936 | 0.000 |
| Yadav *et al*[33] | 32.20 | 1.120 | 1.25 | 30.00 | 34.39 | 28.750 | 0.000 |
| Midha *et al*[17] | 32.80 | 2.350 | 2.40 | 28.10 | 37.49 | 28.725 | 0.000 |
| Gupta[27] | 48.20 | 1.640 | 2.69 | 44.98 | 51.41 | 29.390 | 0.000 |
| Chakraborty *et al*[9] | 17.60 | 1.830 | 3.34 | 14.01 | 21.18 | 9.617 | 0.000 |
| Random effect size | 40.76 | 0.137 | 0.01 | 40.49 | 41.02 | 297.873 | 0.000 |

**B**

Meta Analysis

| Study name | Statistics for each study | | | | | | |
|---|---|---|---|---|---|---|---|
| | Point estimate | SE | Variance | Lower limit | Upper limit | *Z*-value | *P*-value |
| Kusuma *et al*[31] | 21.00 | 1.100 | 1.21 | 18.84 | 23.15 | 19.091 | 0.000 |
| Hazarika *et al*[30] | 33.30 | 0.800 | 0.64 | 31.73 | 34.86 | 41.625 | 0.000 |
| Midha *et al*[17] | 14.50 | 1.500 | 0.25 | 11.10 | 17.90 | 18.600 | 0.000 |
| Todkar *et al*[19] | 7.20 | 0.700 | 0.49 | 5.82 | 8.57 | 10.286 | 0.000 |
| Bhardwaj *et al*[24] | 35.90 | 1.500 | 2.25 | 32.96 | 38.84 | 23.933 | 0.000 |
| Kinra *et al*[32] | 20.00 | 0.900 | 0.81 | 18.23 | 21.76 | 22.222 | 0.000 |
| Rajasekar *et al*[20] | 19.10 | 0.900 | 0.81 | 17.33 | 20.86 | 21.222 | 0.000 |
| Kadu *et al*[21] | 12.75 | 0.700 | 0.49 | 11.37 | 14.12 | 18.214 | 0.000 |
| Bansal *et al*[22] | 32.30 | 1.500 | 2.25 | 29.36 | 35.24 | 21.533 | 0.000 |
| Kaur *et al*[23] | 21.40 | 0.400 | 0.16 | 20.61 | 22.18 | 53.500 | 0.000 |
| Random effect size | 17.91 | 0.223 | 0.05 | 17.47 | 18.34 | 80.259 | 0.000 |

**Figure 2 The overall prevalence rate is represented by the random effect size.** A: Meta-analysis of prevalence of hypertension in the urban Indian population; B: Meta-analysis of prevalence of hypertension in the rural Indian population.

the rural Indian population. Gupta *et al*[8] reported the highest prevalence of hypertension (48.2%) in a recent multi-centric study, conducted in the urban population of India. However, Chakraborty *et al*[9] observed a lower prevalence of hypertension (17.6%) possibly because of the lower age group (18-60 years) and lower socioeconomic strata (slum dwellers) included in the study.

Anand[10] (34.0%), Gupta *et al*[11] (33.4%) and Shanthirani *et al*[12] (21.1%) in the early 2000s, and Reddy *et al*[13] (27.7%), Mohan *et al*[14] (20.0%), Kaur *et al*[15] (27.2%), Prabhakaran *et al*[16] (30.0%) and Midha *et al*[17] (32.8%) in the mid 2000s have observed that the prevalence of hypertension ranged between 20%-40% throughout the decade. However, no significant trend has been observed in the prevalence of hypertension among the studies conducted down the years. The findings of Gupta *et al*[18] reveal a prevalence of 47.9% from Mumbai, Maharashtra, probably because of the stressful lifestyle of subjects in a metropolitan city. The result of the meta-analysis summarizes all these findings and shows the prevalence of hypertension as 40.8% in the urban population of India.

In the rural population, the lowest prevalence (7.2%) was observed by Todkar *et al*[19] in Maharashtra. Midha *et al*[17] (14.5%), Rajasekar *et al*[20] (19.1%) and Kadu *et al*[21] (12.8%) have reported a lower prevalence compared to Bansal *et al*[22] (32.3%) and Kaur *et al*[23] (21.4%). Despite the lack of an obvious trend, the prevalence of hypertension in the rural population is rising swiftly to match up to the urban rates. Bhardwaj *et al*[24] reported the highest prevalence of hypertension (35.9%) in the rural population of Himachal Pradesh. The meta-analysis revealed a prevalence rate of 17.9% among the rural population.

From the available data, it is obvious that the prevalence of hypertension is higher in the urban population of India compared to the rural. No consistent trends are visible with respect to regional variations. In rural populations, the prevalence of hypertension is higher in Himachal Pradesh, while in urban studies prevalence rate is higher in Maharashtra[18,24]. The prevalence rate of hypertension was amongst the highest in metropolitan cities like Mumbai[18]. Moreover, the prevalence of hypertension in rural populations is steadily increasing and is approaching the rates of the urban population. Several studies have reported that there are significant urban-rural differences in metabolic cardiovascular risk factors[25]. Prevalence of smoking is greater in rural men while all other risk factors, such as sedentary lifestyle, obesity, central obesity, hypercholesterolemia, diabetes and the metabolic syndrome, are more common in urban men and women[26]. Most studies on urban-rural differences in cardiovascular risk factors from Haryana, Delhi, Rajasthan and Tamil Nadu have reported greater prevalence of

multiple CVD risk factors in the urban population[27]. As a result, cardiovascular diseases are epidemic in the urban regions of low income countries such as India. However, greater prevalence of cardiovascular risk factors in urban areas in India is in contrast to high income countries where the CVD risk factors are equal in urban and rural areas[28]. Similarly, it has been observed that in the more developed states of India, such as Kerala, the rural-urban differences in cardiometabolic risk factors have largely disappeared and the risk factors are equal or slightly greater in the rural population[29]. Hazarika *et al*[30] reported that even in the rural population of Assam, body mass index and waist-hip ratio were significant risk factors of hypertension. Kusuma *et al*[31] confirmed the hypothesis that acculturation/modernization may elevate the risk of hypertension and that prevalence is generally low among traditional population groups. Kinra *et al*[32] suggested that a nutrition transition (coexistence of over-nutrition and under-nutrition) may have progressed to some parts of rural India. He observed that obesity, dyslipidemia, diabetes and hypertension were more prevalent in higher socioeconomic groups in the rural areas. This epidemiological transition is a cause for serious concern as it is likely that the prevalence of risk factors, and thereby the prevalence of hypertension and cardiovascular diseases, would rise with the socioeconomic development of rural areas.

Yadav *et al*[33] observed that there was a high prevalence of cardiovascular risk factors in the general population [central obesity (86.7%), elevated LDL cholesterol (22.8%), abnormal glucose tolerance (41.6%) and smoking (20.3% of males)]. Two or more of the cardiovascular risk factors were present in a higher proportion of hypertensives (66%, OR = 3.0, $P < 0.0001$) and pre-hypertensives, (56%, OR = 2.0, $P < 0.0001$) compared to normotensive subjects (39%). The current rate of hypertension in the urban areas and the rising trend in the rural population is a warning to institute lifestyle changes in the community in order to put a halt to the increasing rates.

### Challenge ahead

The high prevalence of hypertension in the urban and rural population in India presents a formidable challenge to the Indian health system. In countries like India, the out-of-pocket expenditures incurred for non-communicable diseases (NCDs) like hypertension are high, which hits the impoverished households the most. Medicines for these chronic diseases account for a large portion of expenditure. Therefore, population based prevention strategies have a high impact and are cost-effective as these target lifestyle change. Interventions utilizing the power of public policies for reducing salt, fat, sugar and alcohol intake through regulatory and consumer education approaches; increasing physical activity through sound urban planning and creation of activity-promoting environments; increasing fruit and vegetable intake through appropriate agricultural and pricing mechanisms; and implementing comprehensive tobacco control have the potential to prevent a large proportion of disease events in the whole population[34].

Hypertension is easily diagnosable and treatable with lifestyle modifications and effective medicines. Furthermore, hypertension control provides an entry point to deal with other NCDs as any intervention will help to concomitantly address other NCDs as well. This has been taken into cognizance in the newly launched National Programme for Prevention and Control of Cancer, Diabetes, Cardiovascular Diseases and Stroke (NPCDCS), which has hypertension and diabetes as the main focus areas[34]. National strategies will focus on prevention and health promotion as the key to reduce disease burden. Health education programs that promote exercise, weight reduction, early diagnosis and screening are some of the key interventions that will be promoted at various levels of heath facilities[35]. Under the NPCDCS, the strategy for early diagnosis of chronic NCDs will consist of opportunistic screening of persons above the age of 30 years at the point of primary contact with any health care facility[35]. The NCD clinics mandated under NPCDCS could be leveraged to facilitate guidelines based hypertension management with emphasis on generic drugs and those recommended by the Indian Public Health Standards[34].

In conclusion, the high prevalence of hypertension in the urban population and a rising prevalence in the rural population must raise an alarm for policy makers and health care professionals as this is an area where primordial and primary prevention measures can bring about a substantial reduction in cardiovascular morbidity and mortality in the future.

### Limitations

Given the limited amount of data available on the prevalence of hypertension in India, studies conducted in subjects with different bio-social characteristics have been included in the meta-analysis. Nevertheless, the pooled estimate does provide an overview of the magnitude of the problem of hypertension in the Indian population.

## COMMENTS

### Background

The overall prevalence of hypertension in adults aged 25 and over was around 40% in the world in 2008 and a meta-analysis estimated a prevalence of 10-30.9% in India in 2002. This shift in epidemiological profile presents a unique challenge to India's health system as rates of cardiovascular and metabolic disease like hypertension and diabetes, obesity and cancer rise, tuberculosis, diarrheal disease and water borne illnesses remain widespread. According to a 2012 World Health Organization report, non communicable diseases are responsible for two-thirds of the total morbidity burden and about 53% of total deaths in India. Hypertension provides an entry point to other non-communicable diseases. Therefore, a precise estimate of the prevalence of hypertension in the urban and rural population of the India is required to assess the magnitude of the problem that has to be addressed.

### Research frontiers

Very few studies are available on the prevalence of hypertension in India. A consolidated estimate of hypertension from various studies conducted in different regions of the country can aid the development of preventive strategies. The difference in prevalence between the urban and rural population has also

been studied to provide an insight to the kind of preventive and promotive services required.

### Innovations and breakthroughs

It is possible that an insight into the magnitude of the problem of hypertension can help in shaping the preventive programs and policies specific for the rural and urban population.

### Applications

Very few multi-centric studies are available in India; therefore, the main application of this meta-analysis is to consolidate the available data to determine the burden of hypertension in country.

### Terminology

A meta-analysis integrates the quantitative findings from separate but similar studies and provides a numerical estimate of the overall effect of interest. Different weights are assigned to the different studies for calculating the summary or pooled effect. The weighting is related with the inverse of the standard error (and therefore indirectly to the sample size) reported in the studies. Studies with smaller standard error and larger sample size are given more weight in the calculation of the pooled effect size. The meta-analysis table lists the prevalence of hypertension (expressed as a percentage), with their 95%CI found in the individual studies included in the meta-analysis. The pooled proportion (prevalence) with 95%CI is given for the random effects model. The random effects model will tend to give a more conservative estimate (*i.e.*, with wider confidence interval), but the results are more valid as they take into account any inherent heterogeneity. Under the random effects model, the true effects in the studies are assumed to vary between studies and the summary effect is the weighted average of the effects reported in the different studies.

### Peer review

This manuscript is a meta-analysis on the prevalence of hypertension in India. Its results have provided evidence for policies and interventions for hypertension. It is well written.

## REFERENCES


1 **World Health Organization**. Global Health Repository. Available from: URL: http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence_text/en/index.html. Last accessed April 24, 2013

2 World Health Report-2002. Reducing Risks, Promoting Healthy Life. Chapter 4, p-12. Available from: URL: http://www.who.int/whr/2002/en/whr02_ch4.pdf. Last accessed April 24, 2013

3 Available from: URL: http://www.who.int/world-health-day/en/. Last accessed April 24, 2013

4 **Rodgers A**, Lawes C, MacMahon S. Reducing the global burden of blood pressure-related cardiovascular disease. *J Hypertens Suppl* 2000; **18**: S3-S6 [PMID: 10939783]

5 **Chobanian AV**, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, Jones DW, Materson BJ, Oparil S, Wright JT, Roccella EJ. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA* 2003; **289**: 2560-2572 [PMID: 12748199]

6 **Padmavati S**. Prevention of heart disease in India in the 21st century: need for a concerted effort. *Indian Heart J* 2002; **54**: 99-102 [PMID: 11999100]

7 World Development Report-2006. Equity and Development. A co-publication of the World Bank and Oxford University Press

8 **Gupta R**, Pandey RM, Misra A, Agrawal A, Misra P, Dey S, Rao S, Menon VU, Kamalamma N, Vasantha Devi KP, Revathi K, Vikram NK, Sharma V, Guptha S. High prevalence and low awareness, treatment and control of hypertension in Asian Indian women. *J Hum Hypertens* 2012; **26**: 585-593 [PMID: 21881598 DOI: 10.1038/jhh.2011.79]

9 **Chakraborty R**, Bose K, Kozieł S. Waist circumference in determining obesity and hypertension among 18-60 years old Bengalee Hindu male slum dwellers in Eastern India. *Ann Hum Biol* 2011; **38**: 669-675 [PMID: 21838593]

10 **Anand MP**. Prevalence of hypertension amongst Mumbai executives. *J Assoc Physicians India* 2000; **48**: 1200-1201 [PMID: 11280229]

11 **Gupta R**, Gupta VP, Sarna M, Bhatnagar S, Thanvi J, Sharma V, Singh AK, Gupta JB, Kaul V. Prevalence of coronary heart disease and risk factors in an urban Indian population: Jaipur Heart Watch-2. *Indian Heart J* 2002; **54**: 59-66 [PMID: 11999090]

12 **Shanthirani CS**, Pradeepa R, Deepa R, Premalatha G, Saroja R, Mohan V. Prevalence and risk factors of hypertension in a selected South Indian population--the Chennai Urban Population Study. *J Assoc Physicians India* 2003; **51**: 20-27 [PMID: 12693449]

13 **Reddy KS**, Prabhakaran D, Chaturvedi V, Jeemon P, Thankappan KR, Ramakrishnan L, Mohan BV, Pandav CS, Ahmed FU, Joshi PP, Meera R, Amin RB, Ahuja RC, Das MS, Jaison TM. Methods for establishing a surveillance system for cardiovascular diseases in Indian industrial populations. *Bull World Health Organ* 2006; **84**: 461-469 [PMID: 16799730]

14 **Mohan V**, Deepa M, Farooq S, Datta M, Deepa R. Prevalence, awareness and control of hypertension in Chennai--The Chennai Urban Rural Epidemiology Study (CURES-52). *J Assoc Physicians India* 2007; **55**: 326-332 [PMID: 17844691]

15 **Kaur P**, Rao TV, Sankarasubbaiyan S, Narayanan AM, Ezhil R, Rao SR, Gupte MD. Prevalence and distribution of cardiovascular risk factors in an urban industrial population in south India: a cross-sectional study. *J Assoc Physicians India* 2007; **55**: 771-776 [PMID: 18290552]

16 **Prabhakaran D**, Shah P, Chaturvedi V, Ramakrishnan L, Manhapra A, Reddy KS. Cardiovascular risk factor prevalence among men in a large industry of northern India. *Natl Med J India* 2005; **18**: 59-65 [PMID: 15981439]

17 **Midha T**, Idris MZ, Saran RK, Srivastav AK, Singh SK. Prevalence and determinants of hypertension in the urban and rural population of a north Indian district. *East Afr J Public Health* 2009; **6**: 268-273 [PMID: 20803917]

18 **Gupta PC**, Gupta R, Pednekar MS. Hypertension prevalence and blood pressure trends in 88 653 subjects in Mumbai, India. *J Hum Hypertens* 2004; **18**: 907-910 [PMID: 15306829 DOI: 10.1038/sj.jhh.1001763]

19 **Todkar SS**, Gujarathi VV, Tapare VS. Period prevalence and sociodemographic factors of hypertension in rural maharashtra: a cross-sectional study. *Indian J Community Med* 2009; **34**: 183-187 [PMID: 20049292 DOI: 10.4103/0970-0218.55269]

20 **Rajasekar VD**, Krishnagopal L, Mittal A, Singh Z, Purty AJ, Binu VS, Iillayabharathi V. Prevalence and risk factors for hypertension in a rural area of Tamil Nadu, South India. *Indian J Med Spec* 2012; **3**: 12-17 [DOI: 10.7713/ijms.2012.0004]

21 **Kadu AV**, Mane SS, Lakde RN, Vedpathak VL, Gaikwad AE, Choudhari SG. Prevalence of Hypertension in the Rural Community of Central Maharashtra, India. *Int J Med Public Health* 2012; **2**; 39-45

22 **Bansal SK**, Saxena V, Kandpal SD, Gray WK, Walker RW, Goel D. The prevalence of hypertension and hypertension risk factors in a rural Indian community: A prospective door-to-door study. *J Cardiovasc Dis Res* 2012; **3**: 117-123 [PMID: 22629029 DOI: 10.4103/0975-3583.95365]

23 **Kaur P**, Rao SR, Radhakrishnan E, Rajasekar D, Gupte MD. Prevalence, awareness, treatment, control and risk factors for hypertension in a rural population in South India. *Int J Public Health* 2012; **57**: 87-94 [PMID: 21947549]

24 **Bhardwaj R**, Kandori A, Marwah R, Vaidya P, Singh B, Dhiman P, Sharma A. Prevalence, awareness and control of hypertension in rural communities of Himachal Pradesh. *J Assoc Physicians India* 2010; **58**: 423-424, 429 [PMID: 21121207]

25 **Gupta R**, Joshi P, Mohan V, Reddy KS, Yusuf S. Epidemiology and causation of coronary heart disease and stroke in India. *Heart* 2008; **94**: 16-26 [PMID: 18083949 DOI: 10.1136/hrt.2007.132951]

26 **Gupta R**, Gupta VP. Urban-rural differences in coronary risk factors do not fully explain greater urban coronary heart dis-

ease prevalence. *J Assoc Physicians India* 1997; **45**: 683-686
27 **Gupta R**. Epidemiology and regional variations in cardiovascular disease and risk factors in India. *J Prev Cardiol* 2011; **1**: 7-15
28 **Stuckler D**. Population causes and consequences of leading chronic diseases: a comparative analysis of prevailing explanations. *Milbank Q* 2008; **86**: 273-326 [PMID: 18522614 DOI: 10.1111/j.1468-0009.2008.00522.x]
29 **Thankappan KR**, Shah B, Mathur P, Sarma PS, Srinivas G, Mini GK, Daivadanam M, Soman B, Vasan RS. Risk factor profile for chronic non-communicable diseases: results of a community-based study in Kerala, India. *Indian J Med Res* 2010; **131**: 53-63 [PMID: 20167974]
30 **Hazarika NC**, Narain K, Biswas D, Kalita HC, Mahanta J. Hypertension in the native rural population of Assam. *Natl Med J India* 2004; **17**: 300-304 [PMID: 15736549]
31 **Kusuma YS**, Babu BV, Naidu JM. Prevalence of hypertension in some cross-cultural populations of Visakhapatnam district, South India. *Ethn Dis* 2004; **14**: 250-259 [PMID: 15132211]
32 **Kinra S**, Bowen LJ, Lyngdoh T, Prabhakaran D, Reddy KS, Ramakrishnan L, Gupta R, Bharathi AV, Vaz M, Kurpad AV, Smith GD, Ben-Shlomo Y, Ebrahim S. Sociodemographic patterning of non-communicable disease risk factors in rural India: a cross sectional study. *BMJ* 2010; **341**: c4974 [PMID: 20876148 DOI: 10.1136/bmj.c4974]
33 **Yadav S**, Boddula R, Genitta G, Bhatia V, Bansal B, Kongara S, Julka S, Kumar A, Singh HK, Ramesh V, Bhatia E. Prevalence & risk factors of pre-hypertension & hypertension in an affluent north Indian population. *Indian J Med Res* 2008; **128**: 712-720 [PMID: 19246794]
34 **Mohan S**, Campbell N, Chockalingam A. Time to effectively address hypertension in India. *Indian J Med Res* 2013; **137**: 627-631 [PMID: 23703328]
35 Operational guidelines. National programme for prevention and control of cancer, diabetes, cardiovascular diseases and stroke (NPCDCS). Directorate General of Health Services. Ministry of Health and Family Welfare. Government of India. Available from: URL: http://health.bih.nic.in/Docs/Guidelines/Guidelines-NPCDCS.pdf. Last accessed August 2013

**P- Reviewers** Bergese SD, Tan XR **S- Editor** Wen LL
**L- Editor** Roemmele A **E- Editor** Zheng XM

World Journal of Meta-Analysis

Online Submissions: http://www.wjgnet.com/esps/
wjma@wjgnet.com
doi:10.13105/wjma.v1.i2.90

*World J Meta-Anal* 2013 August 26; 1(2): 90-96
ISSN 2308-3840 (online)


META-ANALYSIS

# Subclinical hypothyroidism and the metabolic syndrome: A meta-analysis of cross-sectional studies

Yi-Cong Ye, Hong-Zhi Xie, Xi-Liang Zhao, Shu-Yang Zhang

Yi-Cong Ye, Hong-Zhi Xie, Xi-Liang Zhao, Shu-Yang Zhang, Department of Cardiology, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing 100730, China

Author contributions: Zhang SY contributed to the concept design, data analysis and manuscript revision; Ye YC and Xie HZ contributed to the database search, study selection, quality assessment, statistic analysis, data interpretation and drafting the manuscript; Zhao XL contributed to data analysis and drafting the manuscript.

Correspondence to: Dr. Shu-Yang Zhang, Department of Cardiology, Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, Dongcheng District Dongdansantiao, Beijing 100730, China. zhangebmg@gmail.com
Telephone: +86-10-69155068 Fax: +86-10-69155068
Received: February 20, 2013 Revised: May 23, 2013
Accepted: June 1, 2013
Published online: August 26, 2013

## Abstract

**AIM:** To determine the relationship between subclinical hypothyroidism (SCH) and the metabolic syndrome (MS).

**METHODS:** We performed a systematic search of databases [MEDLINE (July 1950 to July 2012), EMBASE (July 1966 to July 2012)] and the references of identified studies. Completely published cross-sectional studies of a general population involving SCH and the MS were included. The pooled odds ratio and weighted mean difference (WMD) for the outcomes were calculated using random-effects models.

**RESULTS:** Six cross-sectional studies with 19546 participants were included. In total, 398 of 1324 participants (30.06%) in the SCH group had the MS compared with 4975 of 18222 participants (27.30%) in the euthyroid group [OR = 1.20; 95%CI: 1.05-1.36; $P$ = 0.004; $\chi^2$ = 2.53 ($P$ = 0.773); $I^2$ = 0%]. Further analysis of the components of the MS showed that SCH was associated with increased body mass index (WMD, 0.32 kg/m$^2$; 95%CI: 0.04-0.61; $P$ = 0.026), systolic blood pressure (WMD, 2.62 mmHg; 95%CI: 1.35-3.89; $P$ < 0.001) and triglyceride (WMD, 0.25 mmol/L; 95%CI: 0.23-0.28; $P$ < 0.001).

**CONCLUSION:** Based on the cross-sectional data, SCH may be associated with an increased risk of the MS, which could be attributed to the increased risk of metabolic components.



**Key words:** Subclinical hypothyroidism; Metabolic syndrome; Meta-analysis

**Core tip:** A recent meta-analysis of individual data concluded that subclinical hypothyroidism (SCH) is associated with an increased risk of coronary heart disease (CHD) and CHD mortality. Meanwhile, it has been well recognized that the metabolic syndrome (MS) is associated with increased cardiovascular events and all-cause mortality. Our meta-analysis of cross-sectional data demonstrated that SCH may be associated with an increased risk of the MS, which may explain the relationship between SCH and increased risk of CHD.

Ye YC, Xie HZ, Zhao XL, Zhang SY. Subclinical hypothyroidism and the metabolic syndrome: A meta-analysis of cross-sectional studies. *World J Meta-Anal* 2013; 1(2): 90-96 Available from: URL: http://www.wjgnet.com/2308-3840/full/v1/i2/90.htm DOI: http://dx.doi.org/10.13105/wjma.v1.i2.90

## INTRODUCTION

Subclinical thyroid disease is defined biochemically and subclinical hypothyroidism (SCH) occurs when serum thyroid-stimulating hormone (TSH) concentrations are raised and serum thyroid hormone concentrations are normal[1].

Over the past several decades, a plethora of publications has established that SCH is associated with cardiovascular disease and a meta-analysis of individual data by Rodondi *et al*[2] concluded that SCH is associated with an increased risk of coronary heart disease (CHD) and CHD mortality in those with higher TSH levels, particularly in those with a TSH concentration of 10 mIU/L or greater.

The metabolic syndrome (MS) is a cluster of multiple cardiovascular risk factors, including central obesity, glucose intolerance, diabetes, dyslipidemia and elevated blood pressure[3]. It has been well recognized that the MS is associated with increased cardiovascular events and all-cause mortality[4]. Recently, several cross-sectional studies have investigated a potential relationship between SCH and the MS, with conflicting findings[5-10]. To clarify this issue, we performed a meta-analysis of the published cross-sectional data in a general population to determine the association between SCH and the MS.

## MATERIALS AND METHODS

We designed a protocol that detailed the objective of our analysis, the criteria for study inclusion/exclusion, the assessment of study quality, the primary outcome and the statistical methods in accordance with the meta-analysis of observational studies in epidemiology[11] and PRISMA statements[12].

### Data sources and searches

We conducted a search of MEDLINE (1950 to July 2012) and EMBASE (1966 to July 2012) *via* EMBASE.com to identify all cross-sectional studies of a general population involving SCH and the MS [search strategy: (1) "hypothyroidism"/exp OR hypothyroidism; (2) "thyroid dysfunction"/exp OR thyroid dysfunction; (3) "thyroid disease"/exp OR thyroid disease; (4) "thyrotropin"/exp OR thyrotropin; (5) "thyroid stimulation hormone"/exp OR thyroid stimulation hormone; (6) "MS x"/exp OR MS x; (7) "MS"/exp OR MS; (8) 1 OR 2 OR 3 OR 4 OR 5; (9) 6 OR 7; (10) 8 AND 9], as well as a search of the Cochrane Database for Systemic Reviews. In addition, we performed a manual search of the literature using the references of original manuscripts, reviews and meta-analyses.

### Study selection

Two reviewers (Ye YC and Xie HZ) independently determined the study eligibility. Disagreements were resolved by consensus. The eligibility criteria for study inclusion were: (1) cross-sectional studies; (2) studies of a general population; (3) studies that reported the numbers of participants with the MS from among subjects with SCH and euthyroid; and (4) completely published studies (exclusive of unpublished material and abstracts). No language restriction was imposed. The κ value between two reviewers (Ye YC and Xie HZ) was 0.97 for the first screen (based on title and abstract) and 1.0 for the full-text screen.

### Data extraction

Data extraction was carried out independently by two authors in duplicate (Ye YC and Xie HZ). Disagreements were resolved by discussion between the two reviewing authors. From each included study, information was extracted on: (1) the characteristics of the study population; (2) the criteria for the MS; (3) TSH reference; (4) the numbers of participants with the MS in SCH and euthyroid groups; and (5) the components of the MS in SCH and euthyroid groups [including waist circumference, body mass index (BMI), systolic blood pressure (SBP), diastolic blood pressure (DBP), fasting plasma glucose, triglyceride (TG) and high-density lipoprotein cholesterol (HDL-C)].

### Risk of bias

Risk of bias was described and judged in following domains[13]: (1) Define the source of information (survey, record review); (2) List inclusion and exclusion criteria for exposed and unexposed subjects (cases and controls) or refer to previous publications; (3) Indicate time period used for identifying individuals; (4) Indicate whether or not subjects were consecutively enrolled if not population-based; (5) Indicate if evaluators of subjective components of study were masked to other aspects of the status of the participants; (6) Describe any assessments undertaken for quality assurance purposes (*e.g.*, test/retest of primary outcome measurements); (7) Explain any individual exclusions from analysis; (8) Describe how confounding was assessed and/or controlled; (9) If applicable, explain how missing data were handled in the analysis; and (10) Summarize individual response rates and completeness of data collection. The judgments were assessed independently by two authors (Ye YC and Xie HZ) based on the published report and protocols of the studies. Disagreements were resolved by consensus. The judgments involved the answers "yes" (indicating a low risk of bias), "no" (indicating a high risk of bias), and "unclear" (if risk of bias is unknown or if an entry is not relevant to the study).

### Statistical analysis

The primary outcome was the unadjusted OR for the MS between the SCH and the euthyroid group. The secondary outcomes were the weighted mean difference (WMD) of the components of the MS between the SCH and the euthyroid group. The pooled unadjusted ORs and WMD for the outcomes were calculated using fixed effects models since there may be potential heterogeneity among the studies. The heterogeneity between the results of different studies was examined using $\chi^2$ tests for significance (a $P$-value < 0.10 was considered to be statistically significant) and the inconsistence was examined using $I^2$ tests. Subgroup analysis was used to explore the potential race heterogeneity (Chinese population and non Chinese population). Sensitivity analysis was used to explore the degree to which the main findings of our meta-analysis were affected by individual studies. Publication

**Table 1 Characteristics of included studies**

| Studies | Age (yr) | Male (%) | MS criteria | Sample size | Definition of SCH | Euthyroid | | | SCH | | | Inclusion criteria | Exclusion criteria |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | No. | MS | mTSH | No. | MS | mTSH | | |
| Hergenç *et al*[5] | 51.99 | 46.70 | NCEP /ATPⅢ | 488 | TSH > 4.2 μU/mL | 465 | 193 | 1.12 | 23 | 10 | 7.54 | Age > 34 yr | NA |
| Garduño-Garcia Jde *et al*[6] | 42.3 | 48.80 | NCEP /ATPⅢ | 3033 | TSH: 4.5-10 mIU/L FT4: 10-25 pmol/L | 2771 | 876 | NA | 262 | 84 | NA | Aged 18-70 yr | Thyroid disease, diabetes, cardiovascular disease, cerebral vascular disease, amputations, pregnancy, corticosteroid use, active liver disease, and renal dysfunction |
| Lai *et al*[7] | NA | NA | NCEP /ATPⅢ | 6254 | TSH > 5.0 μU /mL Normal FT4 | 6123 | 1871 | NA | 131 | 43 | 8.5 | Age > 65 yr | Taking medications for control of glucose, blood pressure, dyslipidemia, and thyroid function |
| Lai *et al*[8] | 45.1 | NA | China Diabetes Society | 1385 | TSH > 4.8 mIU/L Normal FT3/FT4 | 1283 | 238 | NA | 102 | 25 | NA | Aged 18-85 yr | History of thyroid disease; Taking medication such as thyroxine/antithyroid drugs, glucocorticoid, antiepileptic and contraceptive drugs; Pregnant/within the first year of postpartum period; Overt hypothyroidism/ hyperthyroidism |
| Liu *et al*[9] | 48.9 | 38.20 | International Diabetes Federation | 6339 | TSH > 4.5 mIU/L Normal FT3/FT4 | 5801 | 1236 | 2 | 538 | 138 | 6.5 | Aged 20-88 yr | Pregnant; Severe renal, liver or heart failure, or abdominal ascites; Taking medicines influencing thyroidal function |
| Waring *et al*[10] | 73.6 | 47 | NCEP /ATPⅢ | 2047 | TSH > 0.35 mIU/L | 1779 | 561 | NA | 268 | 98 | NA | Aged 70-79 yr | Diabetes; Taking medication such as amiodarone, lithium and antithyroid medications |

NCEP/ATPⅢ: National Cholesterol Education Program/Adult Treatment Panel Ⅲ; TSH: Thyroid-stimulating hormone; MS: Metabolic syndrome; SCH: Subclinical hypothyroidism; FT4: Free thyroxine; FT3: Free triiodothyronine; NA: Not available; mTSH: mean/median thyroid-stimulating hormone level.

bias was assessed by the Begg's funnel plot and the Egger weighted regression statistic, with a value of $P < 0.10$ indicating significant publication bias among the included studies[14,15]. All statistical analyses were performed using STATA 11.0 (STATA, TX, United States).

## RESULTS

### Study identification

In total, 515 studies were retrieved from the initial search and 17 studies were reviewed in full text. Six cross-sectional studies investigating a total of 19546 participants were included in the meta-analysis (Figure 1).

### Study characteristics

The analysis included a total of 1324 SCH and 18222 euthyroid is imparticipants. In the individual studies, the sample size ranged from 488 to 6339. The reported mean age of the participants ranged from 42.3 to 73.6 years. Each study included both males and females. Table 1 presents a summary of the characteristics of the study population, the criteria for the MS, the TSH reference and the prevalence of the MS in SCH and euthyroid groups. Of the six studies included, four provided data on the components of the MS in SCH and euthyroid groups (Table 2). The results of quality assessment of included studies (risk of bias) were summarized and are presented in Figure 2.

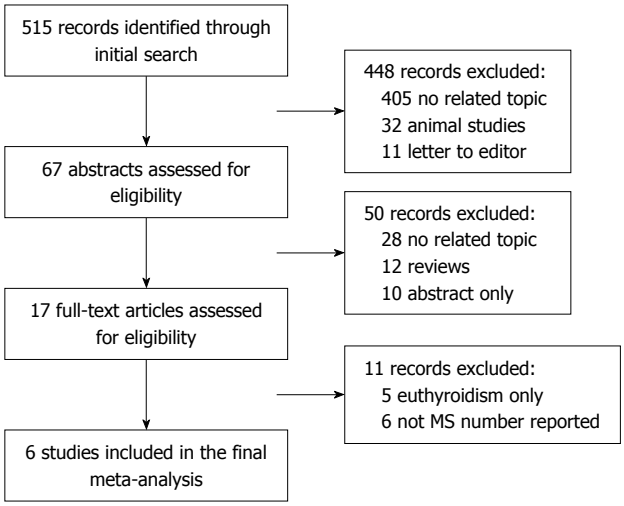



**Figure 1 Flowchart of study selection.** MS: Metabolic syndrome.

### Data analysis of outcomes

In total, 4975 of the 18222 participants (27.30%) in the euthyroid group fulfilled the criteria for the MS, compared to 398 of the 1324 participants (30.06%) in the

**Table 2 Components of the metabolic syndrome in each study**

| Studies | Thyroid function | *n* | Waist (cm) | BMI (kg/m$^2$) | SBP (mmHg) | DBP (mmHg) | FBG (mmol/L) | TG (mmol/L) | HDL-C (mmol/L) |
|---|---|---|---|---|---|---|---|---|---|
| Hergenç *et al*[5] | Euthyroid | 465 | 95.4 ± 10.5 | 29.3 ± 4.5 | 126.5 ± 20.6 | 80.9 ± 10.9 | 5.77 ± 2.30 | 1.92 ± 1.24 | 1.14 ± 0.32 |
| | SCH | 23 | 93.0 ± 10.5 | 29.0 ± 4.2 | 121.6 ± 23.8 | 78.8 ± 13.2 | 5.66 ± 1.51 | 1.81 ± 1.11 | 1.16 ± 0.30 |
| Garduño-Garcia | Euthyroid | 2771 | 94.21 ± 11.18 | 28.69 ± 4.5 | 115 ± 15.12 | 77.29 ± 10.68 | 4.90 ± 1.16 | 2.29 ± 1.73 | 1.10 ± 0.29 |
| Jde *et al*[6] | SCH | 262 | 94.06 ± 11.18 | 28.98 ± 5.1 | 118 ± 16.23 | 78.22 ± 10.87 | 4.91 ± 0.973 | 2.59 ± 4.03 | 1.13 ± 0.29 |
| Lai *et al*[8] | Euthyroid | 1283 | 80.8 ± 10.3 | 24.3 ± 3.6 | 123 ± 18 | 79 ± 11 | 5.22 ± 0.03 | 1.47 ± 0.03 | 1.33 ± 0.37 |
| | SCH | 102 | 80.7 ± 9.6 | 24.5 ± 3.3 | 122 ± 19 | 77 ± 11 | 5.12 ± 0.11 | 1.73 ± 0.12 | 1.26 ± 0.27 |
| Liu *et al*[9] | Euthyroid | 5801 | 80.4 ± 11.1 | 24.4 ± 6.6 | 126.8 ± 19.2 | 81.9 ± 11.3 | 5.2 ± 1.7 | 1.5 ± 1.6 | 1.3 ± 0.4 |
| | SCH | 538 | 81.1 ± 12.4 | 24.8 ± 3.9 | 130.2 ± 20.8 | 83.4 ± 12.0 | 5.3 ± 1.5 | 1.6 ± 1.7 | 1.3 ± 0.6 |

SCH: Subclinical hypothyroidism; BMI: Body mass index; SBP: Systolic blood pressure; DBP: Diastolic blood pressure; FBG: Fasting blood glucose; TG: Triglyceride; HDL-C: High density lipoprotein cholesterol.

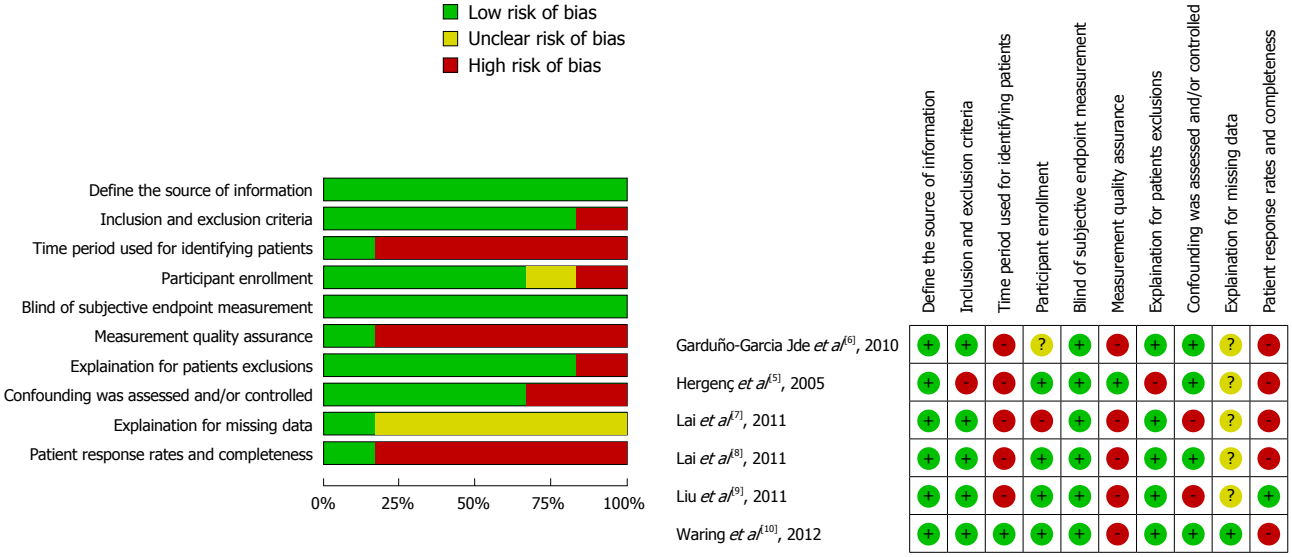



**Figure 2 Risk of bias assessment.**

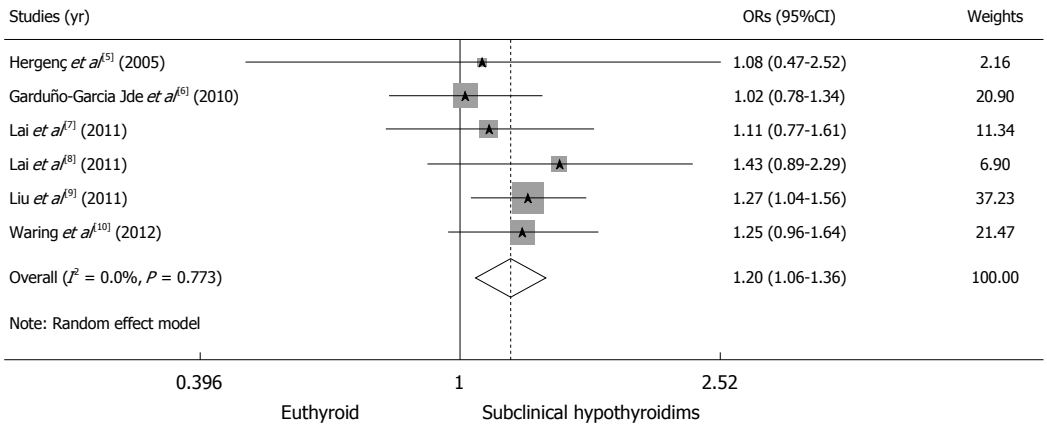



**Figure 3 Subclinical hypothyroidism *vs* euthyroid on pooled ORs for the metabolic syndrome.**

SCH group. The pooled risk of the MS is significantly higher in the SCH group than in the euthyroid group (OR, 1.20; 95%CI: 1.05-1.36; $P = 0.004$). We obtained a $\chi^2$ value of 2.53 ($P = 0.773$) and an $I^2$ value of 0%, indicating the absence of heterogeneity among the studies analyzed (Figure 3).

Further meta-analysis of the components of the MS showed that SCH was associated with increased BMI (WMD, 0.32 kg/m$^2$; 95%CI: 0.04-0.61; $P = 0.026$), SBP (WMD, 2.62 mmHg; 95%CI: 1.35-3.89; $P < 0.001$) and TG (WMD, 0.25 mmol/L; 95%CI: 0.23-0.28; $P < 0.001$). The rest of the components, including waist circumfer-

**Table 3 Sensitivity analysis of the effect of exclusion of individual studies on pooled ORs with 95%CI**

| Studies excluded | Pooled ORs of remaining studies (95%CI) | *P* for OR | $I^2$ (%) | $\chi^2$ | *P* for $\chi^2$ |
|---|---|---|---|---|---|
| Hergenç *et al*[5] | 1.20 (1.06-1.36) | 0.004 | 0 | 2.47 | 0.650 |
| Garduño-Garcia Jde *et al*[6] | 1.25 (1.09-1.44) | 0.002 | 0 | 0.84 | 0.934 |
| Lai *et al*[7] | 1.21 (1.06-1.38) | 0.005 | 0 | 2.34 | 0.673 |
| Lai *et al*[8] | 1.18 (1.04-1.35) | 0.011 | 0 | 1.97 | 0.742 |
| Liu *et al*[9] | 1.15 (0.99-1.35) | 0.071 | 0 | 1.97 | 0.742 |
| Waring *et al*[10] | 1.20 (1.03-1.36) | 0.018 | 0 | 2.40 | 0.663 |

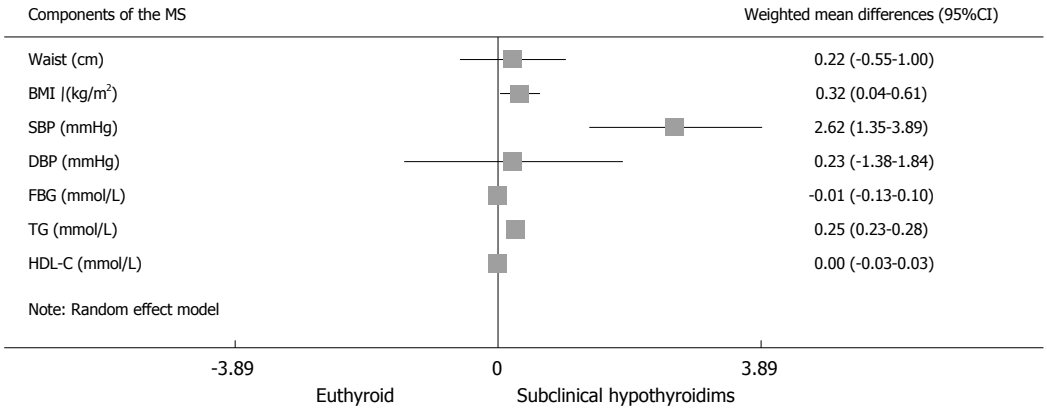



**Figure 4 Weighted mean differences for the components of the metabolic syndrome.** MS: Metabolic syndrome; BMI: Body mass index; SBP: Systolic blood pressure; DBP: Diastolic blood pressure; FBG: Fasting blood glucose; TG: Triglyceride; HDL-C: High-density lipoprotein cholesterol.

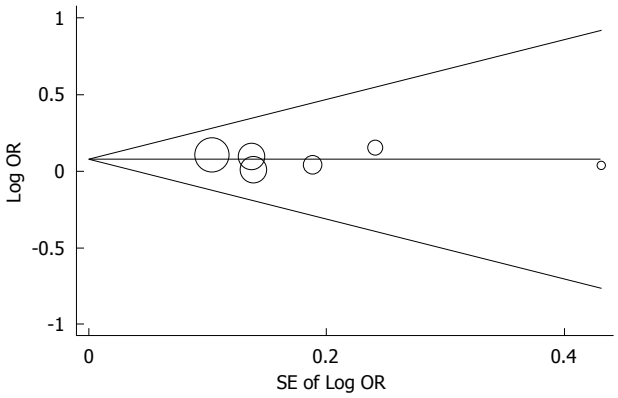



**Figure 5 Begg's funnel plot with pseudo 95%CI for publication bias.**

ence, DBP, fasting blood glucose and HDL-C, were not significantly different between the two groups (Figure 4).

The subgroup analysis indicated that the OR for the MS in Chinese studies did not significantly differ from that in non Chinese studies (OR for MS in Chinese studies: 1.256; 95%CI: 1.063-1.484; OR for MS in non Chinese studies: 1.130; 95%CI: 0.938-1.360, *P* for interaction = 0.404). The sensitivity analysis indicated that the pooled ORs excluding each individual study were comparable (Table 3). No significant publication bias was found after assessment using the Begg's funnel plot (*P* = 0.573) and the Egger weighted regression statistic (*P* = 0.833) (Figure 5).

## DISCUSSION

In this meta-analysis of 19546 participants from six cross-sectional studies of a general population, SCH was associated with an increased risk of the MS compared with being euthyroid. To our best knowledge, this is the first meta-analysis showing the relationship between SCH and the MS. The previous meta-analysis demonstrated a possible association between SCH and increase risk of CHD[2]. This association might be mediated by the effect of these risk factors because the MS is one of the established risk factors for cardiovascular disease[4,16]. However, whether this association between SCH and coronary heart is independent or dependent of MS is still unclear and further investigation is needed.

Even in euthyroid populations, the relationship between thyroid function and the MS has been well investigated, with results consistent with those of our meta-analysis. In a cross-sectional study in the Netherlands, free T4 levels within the normal reference range were significantly related to four of the five components of the MS (*P* for waist circumference = 0.038, *P* for TGs = 0.023, *P* for HDL-C = 0.007 and *P* for SBP = 0.019), independent of insulin resistance[17]. Park *et al*[18] found a close relationship between TSH and the MS in euthyroid postmenopausal women (adjusted OR for the MS 1.55; 95%CI: 1.26-1.89; *P* < 0.001).

It was previously reported that overt hypothyroidism was associated with the MS[19], presumably because of the associated hypercholesterolemia and hypertension[20]. The reason why SCH is associated with the MS could be also attributed to the increased risk of the components of the MS. In a study evaluating 27097 individuals > 40 years of age without a diagnosis of thyroid disease, serum TSH, even within the normal range, was positively associated

with BMI (*P* for trend in all BMI groups < 0.001)[21]. The relationship between blood pressure and thyroid function has also been well investigated but with conflicting findings[22-25]. A recent meta-analysis of seven cross-sectional studies concluded that SCH is associated with increased blood pressure, whereas subclinical hyperthyroidism is not[26]. Hypothyroidism has been associated with dyslipidemia, which is characterized by increased levels of total cholesterol and low-density lipoprotein cholesterol[27]. In the Colorado thyroid disease prevalence study, the total cholesterol, low-density lipoprotein cholesterol and TG levels of SCH subjects were significantly greater than the corresponding lipid levels in euthyroid subjects ($P < 0.001$ for total cholesterol and low-density lipoprotein cholesterol; $P = 0.02$ for TG), whereas no difference was found in HDL-C level[28]. Our meta-analysis of metabolic components has further confirmed these conclusions.

Another important issue is whether SCH should be treated with thyroid hormones or whether hormone replacement therapy could reduce the cardiovascular risk and mortality in SCH. A systematic review of randomized controlled trials showed that thyroid hormone replacement in SCH could improve the lipid profile (WMD for low-density lipoprotein cholesterol: -10.77 mg/dL; 95%CI: -21.57-0.04; $P = 0.051$)[29] and recent data from a cohort study indicated that treatment of SCH with levothyroxine was associated with less ischemic heart disease in younger individuals compared to those without treatment (adjusted hazard ratio 0.61; 95%CI: 0.39-0.95)[30]. Since our meta-analysis indicated a higher risk of the MS in SCH, screening and treatment of these metabolic components are necessary to further reduce the risk of cardiovascular events.

It is notable that the included study from Waring *et al*[10] reported not only cross-sectional data but also prospective data, making it the only prospective study on SCH and the MS in the literature. After 6 years of follow-up, the risk of the MS in SCH individuals with a TSH > 10 mIU/L was similar to that in the baseline cross-sectional analysis but with a wider confidence interval (OR for cross-sectional analysis, 2.3; 95%CI: 1.0-5.0; OR for prospective analysis, 2.2; 95%CI: 0.6-7.5). However, this study seemed to be underpowered because a very small number of participants with marked SCH were included[10].

Our study has several limitations. Firstly, although our heterogeneity test showed no statistically significant heterogeneity among studies, it is still impossible that our final results would be confounded by the different definitions of SCH and MS among the included studies, as well as the different inclusion/exclusion criteria. Secondly, this is a study-level and we used unadjusted OR instead of adjusted OR; thus, it is impossible to exclude the confounding effects of age, sex and race. Finally, although no statistically significant publication bias was detected in our analysis, we believe there may be data relevant to this topic that have never been published.

Our meta-analysis indicated that SCH may be associated with an increased risk of the MS, which is attributed to the increased risk of metabolic components. A large prospective cohort study is needed to further confirm the conclusion of our study.

## COMMENTS

### Background

Over the past several decades, a plethora of publications has established that subclinical hypothyroidism (SCH) is associated with cardiovascular disease. Meanwhile, it has been well recognized that the metabolic syndrome (MS) is associated with increased cardiovascular events and all-cause mortality.

### Research frontiers

Is SCH associated with a higher risk of the MS?

### Innovations and breakthroughs

This meta-analysis of 6 cross-sectional studies with 19546 participants indicated that SCH may be associated with increased risk of the MS, which is attributed to the increased risk of metabolic components.

### Applications

It may be necessary to screen for the MS in SCH individuals in clinical practice.

### Terminology

SCH occurs when thyroid-stimulating hormone levels are elevated but thyroxine (T4) and triiodothyronine (T3) levels are normal. MS is a name for a group of risk factors that occur together and increase the risk for coronary artery disease, stroke and type 2 diabetes.

### Peer review

This paper reports the results of a meta-analysis of observational studies evaluating the association between SCH and MS and its individual components. Overall methodology is correct and the results are very interesting and valuable.

## REFERENCES


1 **Cooper DS**, Biondi B. Subclinical thyroid disease. *Lancet* 2012; **379**: 1142-1154 [PMID: 22273398 DOI: 10.1016/S0140-6736(11)60276-6]

2 **Rodondi N**, den Elzen WP, Bauer DC, Cappola AR, Razvi S, Walsh JP, Asvold BO, Iervasi G, Imaizumi M, Collet TH, Bremner A, Maisonneuve P, Sgarbi JA, Khaw KT, Vanderpump MP, Newman AB, Cornuz J, Franklyn JA, Westendorp RG, Vittinghoff E, Gussekloo J. Subclinical hypothyroidism and the risk of coronary heart disease and mortality. *JAMA* 2010; **304**: 1365-1374 [PMID: 20858880 DOI: 10.1001/jama.2010.1361]

3 **Eckel RH**, Alberti KG, Grundy SM, Zimmet PZ. The metabolic syndrome. *Lancet* 2010; **375**: 181-183 [PMID: 20109902 DOI: 10.1016/S0140-6736(09)61794-3]

4 **Mottillo S**, Filion KB, Genest J, Joseph L, Pilote L, Poirier P, Rinfret S, Schiffrin EL, Eisenberg MJ. The metabolic syndrome and cardiovascular risk a systematic review and meta-analysis. *J Am Coll Cardiol* 2010; **56**: 1113-1132 [PMID: 20863953 DOI: 10.1016/j.jacc.2010.05.034]

5 **Hergenç G**, Onat A, Albayrak S, Karabulut A, Türkmen S, Sari I, Can G. TSH levels in Turkish adults: Prevalences and associations with serum lipids, coronary heart disease and metabolic syndrome. *Turk J Med Sci* 2005; **35**: 297-304

6 **Garduño-Garcia Jde J**, Alvirde-Garcia U, López-Carrasco G, Padilla Mendoza ME, Mehta R, Arellano-Campos O, Choza R, Sauque L, Garay-Sevilla ME, Malacara JM, Gomez-Perez FJ, Aguilar-Salinas CA. TSH and free thyroxine concentrations are associated with differing metabolic markers in euthyroid subjects. *Eur J Endocrinol* 2010; **163**: 273-278 [PMID: 20516204 DOI: 10.1530/EJE-10-0312]

7 **Lai CC**, Tang SH, Pei D, Wang CY, Chen YL, Wu CZ, Hsiao FC, Chen HS, Wang JY. The prevalence of subclinical thyroid dysfunction and its association with metabolic syndrome in Taiwanese elderly. *Int J Gerontol* 2011; **5**: 25-29

8 **Lai Y**, Wang J, Jiang F, Wang B, Chen Y, Li M, Liu H, Li C, Xue H, Li N, Yu J, Shi L, Bai X, Hou X, Zhu L, Lu L, Wang S,

Xing Q, Teng X, Teng W, Shan Z. The relationship between serum thyrotropin and components of metabolic syndrome. *Endocr J* 2011; **58**: 23-30 [PMID: 21135510 DOI: 10.1507/endocrj.K10E-272]

9 **Liu C**, Scherbaum WA, Schott M, Schinner S. Subclinical hypothyroidism and the prevalence of the metabolic syndrome. *Horm Metab Res* 2011; **43**: 417-421 [PMID: 21512964 DOI: 10.1055/s-0031-1275719]

10 **Waring AC**, Rodondi N, Harrison S, Kanaya AM, Simonsick EM, Miljkovic I, Satterfield S, Newman AB, Bauer DC. Thyroid function and prevalent and incident metabolic syndrome in older adults: the Health, Ageing and Body Composition Study. *Clin Endocrinol* (Oxf) 2012; **76**: 911-918 [PMID: 22187968 DOI: 10.1111/j.1365-2265.2011.04328.x]

11 **Stroup DF**, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; **283**: 2008-2012 [PMID: 10789670 DOI: 10.1001/jama.283.15.2008]

12 **Liberati A**, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med* 2009; **151**: W65-W94 [PMID: 19622512 DOI: 10.7326/0003-4819-151-4-200908180-00136]

13 **Rostom A**, Dubé C, Cranney A. Celiac Disease. Rockville, MD: Agency for Healthcare Research and Quality (US), 2004. (Evidence Reports/Technology Assessments, No. 104.) Appendix D. Quality Assessment Forms. Available from: URL: http://www.ncbi.nlm.nih.gov/books/NBK35156/

14 **Begg CB**, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994; **50**: 1088-1101 [PMID: 7786990 DOI: 10.2307/2533446]

15 **Egger M**, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629-634 [PMID: 9310563 DOI: 10.1136/bmj.315.7109.629]

16 **Ford ES**. Risks for all-cause mortality, cardiovascular disease, and diabetes associated with the metabolic syndrome: a summary of the evidence. *Diabetes Care* 2005; **28**: 1769-1778 [PMID: 15983333 DOI: 10.2337/diacare.28.7.1769]

17 **Roos A**, Bakker SJ, Links TP, Gans RO, Wolffenbuttel BH. Thyroid function is associated with components of the metabolic syndrome in euthyroid subjects. *J Clin Endocrinol Metab* 2007; **92**: 491-496 [PMID: 17090642 DOI: 10.1210/jc.2006-1718]

18 **Park HT**, Cho GJ, Ahn KH, Shin JH, Hong SC, Kim T, Hur JY, Kim YT, Lee KW, Kim SH. Thyroid stimulating hormone is associated with metabolic syndrome in euthyroid postmenopausal women. *Maturitas* 2009; **62**: 301-305 [PMID: 19250778 DOI: 10.1016/j.maturitas.2009.01.007]

19 **Erdogan M**, Canataroglu A, Ganidagli S, Kulaksızoglu M. Metabolic syndrome prevalence in subclinic and overt hypothyroid patients and the relation among metabolic syndrome parameters. *J Endocrinol Invest* 2011; **34**: 488-492 [PMID: 20651468]

20 **Klein I**, Ojamaa K. Thyroid hormone and the cardiovascular system. *N Engl J Med* 2001; **344**: 501-509 [PMID: 11172193 DOI: 10.1056/NEJM200102153440707]

21 **Asvold BO**, Bjøro T, Vatten LJ. Association of serum TSH with high body mass differs between smokers and never-smokers. *J Clin Endocrinol Metab* 2009; **94**: 5023-5027 [PMID: 19846737 DOI: 10.1210/jc.2009-1180]

22 **Iqbal A**, Figenschau Y, Jorde R. Blood pressure in relation to serum thyrotropin: The Tromsø study. *J Hum Hypertens* 2006; **20**: 932-936 [PMID: 17024137 DOI: 10.1038/sj.jhh.1002091]

23 **Duan Y**, Peng W, Wang X, Tang W, Liu X, Xu S, Mao X, Feng S, Feng Y, Qin Y, Xu K, Liu C, Liu C. Community-based study of the association of subclinical thyroid dysfunction with blood pressure. *Endocrine* 2009; **35**: 136-142 [PMID: 19130316 DOI: 10.1007/s12020-008-9138-y]

24 **Chen H**, Xi Q, Zhang H, Song B, Liu X, Mao X, Li J, Shen H, Tang W, Zhang J, Wang Z, Duan Y, Liu C. Investigation of thyroid function and blood pressure in school-aged subjects without overt thyroid disease. *Endocrine* 2012; **41**: 122-129 [PMID: 21986920 DOI: 10.1007/s12020-011-9517-7]

25 **Walsh JP**, Bremner AP, Bulsara MK, O'Leary P, Leedman PJ, Feddema P, Michelangeli V. Subclinical thyroid dysfunction and blood pressure: a community-based study. *Clin Endocrinol* (Oxf) 2006; **65**: 486-491 [PMID: 16984241 DOI: 10.1111/j.1365-2265.2006.02619.x]

26 **Cai Y**, Ren Y, Shi J. Blood pressure levels in patients with subclinical thyroid dysfunction: a meta-analysis of cross-sectional data. *Hypertens Res* 2011; **34**: 1098-1105 [PMID: 21796125 DOI: 10.1038/hr.2011.91]

27 **Duntas LH**. Thyroid disease and lipids. *Thyroid* 2002; **12**: 287-293 [PMID: 12034052 DOI: 10.1089/10507250252949405]

28 **Canaris GJ**, Manowitz NR, Mayor G, Ridgway EC. The Colorado thyroid disease prevalence study. *Arch Intern Med* 2000; **160**: 526-534 [PMID: 10695693 DOI: 10.1001/archinte.160.4.526]

29 **Villar HC**, Saconato H, Valente O, Atallah AN. Thyroid hormone replacement for subclinical hypothyroidism. *Cochrane Database Syst Rev* 2007; CD003419 [PMID: 17636722]

30 **Razvi S**, Weaver JU, Butler TJ, Pearce SH. Levothyroxine treatment of subclinical hypothyroidism, fatal and nonfatal cardiovascular events, and mortality. *Arch Intern Med* 2012; **172**: 811-817 [PMID: 22529180 DOI: 10.1001/archinternmed.2012.1159]

**P- Reviewers** Chao M, Cheung CL, Mariani J, Singh-Franco D, Treglia G **S- Editor** Wen LL **L- Editor** Roemmele A **E- Editor** Zheng XM

百世登
Baishideng®

Published by **Baishideng Publishing Group Co., Limited**
Flat C, 23/F., Lucky Plaza,
315-321 Lockhart Road, Wan Chai, Hong Kong, China
Fax: +852-65557188
Telephone: +852-31779906
E-mail: bpgoffice@wjgnet.com
http://www.wjgnet.com

World Journal of Meta-Analysis

Online Submissions: http://www.wjgnet.com/esps/
wjma@wjgnet.com
www.wjgnet.com

*World J Meta-Anal* 2013 August 26; 1(2): I-V
ISSN 2308-3840 (online)


*INSTRUCTIONS TO AUTHORS*

## GENERAL INFORMATION

*World Journal of Meta-Analysis* (*World J Meta-Anal*, *WJMA*, online ISSN 2308-3840, DOI: 10.13105) is a peer-reviewed open access (OA) academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians, with a specific focus on meta-analysis, systematic review, mixed-treatment comparison, meta-regression, overview of reviews.

### *Aim and scope*

*WJMA* covers a variety of clinical medical fields including allergy, anesthesiology, cardiac medicine, clinical genetics, clinical neurology, critical care, dentistry, dermatology, emergency medicine, endocrinology, family medicine, gastroenterology and hepatology, geriatrics and gerontology, hematology, immunology, infectious diseases, internal medicine, obstetrics and gynecology, oncology, ophthalmology, orthopedics, otolaryngology, pathology, pediatrics, peripheral vascular disease, psychiatry, radiology, rehabilitation, respiratory medicine, rheumatology, surgery, toxicology, transplantation, and urology and nephrology, while maintaining its unique dedication to systematic reviews and meta-analyses.

We encourage authors to submit their manuscripts to *WJMA*. We will give priority to manuscripts that are supported by major national and international foundations and those that are of great basic and clinical significance.

*WJMA* is edited and published by Baishideng Publishing Group (BPG). BPG has a strong professional editorial team composed of science editors, language editors and electronic editors. BPG currently publishes 43 OA clinical medical journals, including 42 in English, has a total of 15471 editorial board members or peer reviewers, and is a world first-class publisher.

### *Columns*

The columns in the issues of *WJMA* will include: (1) Editorial: The editorial board members are invited to make comments on an important topic in their field in terms of its current research status and future directions to lead the development of this discipline; (2) Frontier: The editorial board members are invited to select a highly cited cutting-edge original paper of his/her own to summarize major findings, the problems that have been resolved and remain to be resolved, and future research directions to help readers understand his/her important academic point of view and future research directions in the field; (3) Field of Vision: The editorial board members are invited to write commentaries on classic articles, hot topic articles, or latest articles to keep readers at the forefront of research and increase their levels of clinical research. Classic articles refer to papers that are included in Web of Knowledge and have received a large number of citations (ranking in the top 1%) after being published for more than 2 years, reflecting the quality and impact of papers. Hot topic articles refer to papers that are included in Web of Knowledge and have received a large number of citations after being published for no more than 2 years, reflecting cutting-edge trends in scientific research. Latest articles refer to the latest published high-quality papers that are included in PubMed, reflecting the latest research trends. These commentary articles should focus on the status quo of research, the most important research topics, the problems that have now been resolved and remain to be resolved, and future research directions. Basic information about the article to be commented will be provided as well (including authors, article title, journal name, year, volume, and inclusive page numbers; (4) Minireviews: The editorial board members are invited to write short reviews on recent advances and trends in research to provide readers; (5) Review: To make a systematic review to focus on the status quo of research, the most important research topics, the problems that have now been resolved and remain to be resolved, and future research directions; (6) Topic Highlight: The editorial board members are invited to write a series of articles (7-10 articles) to comment and discuss a hot topic; (7) Meta-Analysis: Covers the systematic review, mixed-treatment comparison, meta-regression, and overview of reviews. To summarize a given quantitative effect, *e.g.*, the clinical effectiveness and safety of clinical treatments by combining data from two or more randomized controlled trials, thereby providing more precise and externally valid estimates than those which would stem from each individual dataset if analyzed separately from the others; (8) Medical Ethics: The editorial board members are invited to write articles about medical ethics to increase readers' knowledge of medical ethics. The topic covers international ethics guidelines, animal studies, clinical trials, organ transplantation, *etc.*; (9) Letters to the Editor: To discuss and make reply to the contributions published in *WJMA*, or to introduce and comment on a controversial issue of general interest; (10) Book Reviews: To introduce and comment on quality monographs; and (11) Autobiography: The editorial board members are invited to write their autobiography to provide readers with stories of success or failure in their scientific research career. The topic covers their basic personal information and information about when they started doing research work, where and how they did research work, what they have achieved, and their lessons from success or failure.

### *Name of journal*

*World Journal of Meta-Analysis*

### *ISSN*

ISSN 2308-3840 (online)

### *Launch date*

May 26, 2013

### *Frequency*

Quarterly

### *Editors-in-Chief*



### *Editorial office*

Room 903, Building D, Ocean International Center,
No. 62 Dongsihuan Zhonglu, Chaoyang District,
Beijing 100025, China
Telephone: +86-10-85381891
Fax: +86-10-85381893
E-mail: wjma@wjgnet.com
http://www.wjgnet.com

***Publisher***
Baishideng Publishing Group Co., Limited
Flat C, 23/F., Lucky Plaza, 315-321 Lockhart Road,
Wan Chai, Hong Kong, China
Telephone: +852-58042046
Fax: +852-31158812
E-mail: bpgoffice@wjgnet.com
http://www.wjgnet.com

***Production center***
Beijing Baishideng BioMed Scientific Co., Limited
Room 903, Building D, Ocean International Center,
No. 62 Dongsihuan Zhonglu, Chaoyang District,
Beijing 100025, China
Telephone: +86-10-85381892
Fax: +86-10-85381893

***BPG US INC***
Baishideng Publishing Group Inc
8226 Regency Drive, Pleasanton,
CA 94588, United States

***Instructions to authors***
Full instructions are available online at http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm.

***Indexed and Abstracted in***
Digital Object Identifier.

## SPECIAL STATEMENT

All articles published in this journal represent the viewpoints of the authors except where indicated otherwise.

***Biostatistical editing***
Statistical review is performed after peer review. We invite an expert in Biomedical Statistics to evaluate the statistical method used in the paper, including *t*-test (group or paired comparisons), chi-squared test, Ridit, probit, logit, regression (linear, curvilinear, or stepwise), correlation, analysis of variance, analysis of covariance, *etc.* The reviewing points include: (1) Statistical methods should be described when they are used to verify the results; (2) Whether the statistical techniques are suitable or correct; (3) Only homogeneous data can be averaged. Standard deviations are preferred to standard errors. Give the number of observations and subjects (*n*). Losses in observations, such as drop-outs from the study should be reported; (4) Values such as ED50, LD50, IC50 should have their 95% confidence limits calculated and compared by weighted probit analysis (Bliss and Finney); and (5) The word 'significantly' should be replaced by its synonyms (if it indicates extent) or the *P* value (if it indicates statistical significance).

***Conflict-of-interest statement***
In the interests of transparency and to help reviewers assess any potential bias, *WJMA* requires authors of all papers to declare any competing commercial, personal, political, intellectual, or religious interests in relation to the submitted work. Referees are also asked to indicate any potential conflict they might have reviewing a particular paper. Before submitting, authors are suggested to read "Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Ethical Considerations in the Conduct and Reporting of Research: Conflicts of Interest" from International Committee of Medical Journal Editors (ICMJE), which is available at: http://www.icmje.org/ethical_4conflicts.html.

Sample wording: [Name of individual] has received fees for serving as a speaker, a consultant and an advisory board member for [names of organizations], and has received research funding from [names of organization]. [Name of individual] is an employee of [name of organization]. [Name of individual] owns stocks and shares in [name of organization]. [Name of individual] owns patent [patent identification and brief description].

***Statement of informed consent***
Manuscripts should contain a statement to the effect that all human studies have been reviewed by the appropriate ethics committee or it should be stated clearly in the text that all persons gave their informed consent prior to their inclusion in the study. Details that might disclose the identity of the subjects under study should be omitted. Authors should also draw attention to the Code of Ethics of the World Medical Association (Declaration of Helsinki, 1964, as revised in 2004).

***Statement of human and animal rights***
When reporting the results from experiments, authors should follow the highest standards and the trial should conform to Good Clinical Practice (for example, US Food and Drug Administration Good Clinical Practice in FDA-Regulated Clinical Trials; UK Medicines Research Council Guidelines for Good Clinical Practice in Clinical Trials) and/or the World Medical Association Declaration of Helsinki. Generally, we suggest authors follow the lead investigator's national standard. If doubt exists whether the research was conducted in accordance with the above standards, the authors must explain the rationale for their approach and demonstrate that the institutional review body explicitly approved the doubtful aspects of the study.

Before submitting, authors should make their study approved by the relevant research ethics committee or institutional review board. If human participants were involved, manuscripts must be accompanied by a statement that the experiments were undertaken with the understanding and appropriate informed consent of each. Any personal item or information will not be published without explicit consents from the involved patients. If experimental animals were used, the materials and methods (experimental procedures) section must clearly indicate that appropriate measures were taken to minimize pain or discomfort, and details of animal care should be provided.

## SUBMISSION OF MANUSCRIPTS

Manuscripts should be typed in 1.5 line spacing and 12 pt. Book Antiqua with ample margins. Number all pages consecutively, and start each of the following sections on a new page: Title Page, Abstract, Introduction, Materials and Methods, Results, Discussion, Acknowledgements, References, Tables, Figures, and Figure Legends. Neither the editors nor the publisher are responsible for the opinions expressed by contributors. Manuscripts formally accepted for publication become the permanent property of Baishideng Publishing Group Co., Limited, and may not be reproduced by any means, in whole or in part, without the written permission of both the authors and the publisher. We reserve the right to copy-edit and put onto our website accepted manuscripts. Authors should follow the relevant guidelines for the care and use of laboratory animals of their institution or national animal welfare committee. For the sake of transparency in regard to the performance and reporting of clinical trials, we endorse the policy of the ICMJE to refuse to publish papers on clinical trial results if the trial was not recorded in a publicly-accessible registry at its outset. The only register now available, to our knowledge, is http://www.clinicaltrials.gov sponsored by the United States National Library of Medicine and we encourage all potential contributors to register with it. However, in the case that other registers become available you will be duly notified. A letter of recommendation from each author's organization should be provided with the contributed article to ensure the privacy and secrecy of research is protected.

Authors should retain one copy of the text, tables, photographs and illustrations because rejected manuscripts will not be returned to the author(s) and the editors will not be responsible for loss or damage to photographs and illustrations sustained during mailing.

***Online submissions***
Manuscripts should be submitted through the Online Submission System at: http://www.wjgnet.com/esps/. Authors are highly recommended to consult the ONLINE INSTRUCTIONS TO AUTHORS (http://www.wjgnet.com/2308-3840/g_info_20100722180909.htm) before attempting to submit online. For assistance, authors encountering problems with the Online Submission System may send an email describing the problem to wjma@wjgnet.com, or by telephone: +86-10-85381892. If you submit your manuscript online, do not make a postal contribution. Repeated online submission for the same manuscript is strictly prohibited.

## MANUSCRIPT PREPARATION

All contributions should be written in English. All articles must be submitted using word-processing software. All submissions must be typed in 1.5 line spacing and 12 pt. Book Antiqua with ample margins. Style should conform to our house format. Required information for each of the manuscript sections is as follows:

***Title page***
**Title:** Title should be less than 12 words.

**Running title:** A short running title of less than 6 words should be provided.

**Authorship:** Authorship credit should be in accordance with the standard proposed by ICMJE, based on (1) substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; (2) drafting the article or revising it critically for important intellectual content; and (3) final approval of the version to be published. Authors should meet conditions 1, 2, and 3.

**Institution:** Author names should be given first, then the complete name of institution, city, province and postcode. For example, Xu-Chen Zhang, Li-Xin Mei, Department of Pathology, Chengde Medical College, Chengde 067000, Hebei Province, China. One author may be represented from two institutions, for example, George Sgourakis, Department of General, Visceral, and Transplantation Surgery, Essen 45122, Germany; George Sgourakis, 2nd Surgical Department, Korgialenio-Benakio Red Cross Hospital, Athens 15451, Greece

**Author contributions:** The format of this section should be: Author contributions: Wang CL and Liang L contributed equally to this work; Wang CL, Liang L, Fu JF, Zou CC, Hong F and Wu XM designed the research; Wang CL, Zou CC, Hong F and Wu XM performed the research; Xue JZ and Lu JR contributed new reagents/analytic tools; Wang CL, Liang L and Fu JF analyzed the data; and Wang CL, Liang L and Fu JF wrote the paper.

**Supportive foundations:** The complete name and number of supportive foundations should be provided, *e.g.*, Supported by National Natural Science Foundation of China, No. 30224801

**Correspondence to:** Only one corresponding address should be provided. Author names should be given first, then author title, affiliation, the complete name of institution, city, postcode, province, country, and email. All the letters in the email should be in lower case. A space interval should be inserted between country name and email address. For example, Montgomery Bissell, MD, Professor of Medicine, Chief, Liver Center, Gastroenterology Division, University of California, Box 0538, San Francisco, CA 94143, United States. montgomery.bissell@ucsf.edu

**Telephone and fax:** Telephone and fax should consist of +, country number, district number and telephone or fax number, *e.g.*, Telephone: +86-10-85381892 Fax: +86-10-85381893

**Peer reviewers:** All articles received are subject to peer review. Normally, three experts are invited for each article. Decision on acceptance is made only when at least two experts recommend publication of an article. All peer-reviewers are acknowledged on Express Submission and Peer-review System website.

***Abstract***
There are unstructured abstracts (no less than 200 words) and structured abstracts. The specific requirements for structured abstracts are as follows:

An informative, structured abstract should accompany each manuscript. Abstracts of original contributions should be structured into the following sections: AIM (no more than 20 words; Only the purpose of the study should be included. Please write the Aim in the form of "To investigate/study/…"), METHODS (no less than 140 words for Original Articles; and no less than 80 words for Brief Articles), RESULTS (no less than 150 words for Original Articles and no less than 120 words for Brief Articles; You should present *P* values where appropriate and must provide relevant data to illustrate how they were obtained, *e.g.*, $6.92 \pm 3.86$ *vs* $3.61 \pm 1.67$, $P < 0.001$), and CONCLUSION (no more than 26 words).

***Key words***
Please list 5-10 key words, selected mainly from *Index Medicus*, which reflect the content of the study.

***Core tip***
Please write a summary of less than 100 words to outline the most innovative and important arguments and core contents in your paper to attract readers.

***Text***
For articles of these sections, original articles and brief articles, the main text should be structured into the following sections: INTRODUCTION, MATERIALS AND METHODS, RESULTS and DISCUSSION, and should include appropriate Figures and Tables. Data should be presented in the main text or in Figures and Tables, but not in both.

***Illustrations***
Figures should be numbered as 1, 2, 3, *etc.*, and mentioned clearly in the main text. Provide a brief title for each figure on a separate page. Detailed legends should not be provided under the figures. This part should be added into the text where the figures are applicable. Keeping all elements compiled is necessary in line-art image. Scale bars should be used rather than magnification factors, with the length of the bar defined in the legend rather than on the bar itself. File names should identify the figure and panel. Avoid layering type directly over shaded or textured areas. Please use uniform legends for the same subjects. For example: Figure 1 Pathological changes in atrophic gastritis after treatment. A: ...; B: ...; C: ...; D: ...; E: ...; F: ...; G: …*etc.* It is our principle to publish high resolution-figures for the E-versions.

***Tables***
Three-line tables should be numbered 1, 2, 3, *etc.*, and mentioned clearly in the main text. Provide a brief title for each table. Detailed legends should not be included under tables, but rather added into the text where applicable. The information should complement, but not duplicate the text. Use one horizontal line under the title, a second under column heads, and a third below the Table, above any footnotes. Vertical and italic lines should be omitted.

***Notes in tables and illustrations***
Data that are not statistically significant should not be noted. $^{a}P < 0.05$, $^{b}P < 0.01$ should be noted ($P > 0.05$ should not be noted). If there are other series of *P* values, $^{c}P < 0.05$ and $^{d}P < 0.01$ are used. A third series of *P* values can be expressed as $^{e}P < 0.05$ and $^{f}P < 0.01$. Other notes in tables or under illustrations should be expressed as $^{1}$F, $^{2}$F, $^{3}$F; or sometimes as other symbols with a superscript (Arabic numerals) in the upper left corner. In a multi-curve illustration, each curve should be labeled with ●, ○, ■, □, ▲, △, *etc.*, in a certain sequence.

***Acknowledgments***
Brief acknowledgments of persons who have made genuine contributions to the manuscript and who endorse the data and conclusions should be included. Authors are responsible for obtaining written permission to use any copyrighted text and/or illustrations.

## REFERENCES

### *Coding system*

The author should number the references in Arabic numerals according to the citation order in the text. Put reference numbers in square brackets in superscript at the end of citation content or after the cited author's name. For citation content which is part of the narration, the coding number and square brackets should be typeset normally. For example, "Crohn's disease (CD) is associated with increased intestinal permeability[1,2]". If references are cited directly in the text, they should be put together within the text, for example, "From references[19,22-24], we know that..."

When the authors write the references, please ensure that the order in text is the same as in the references section, and also ensure the spelling accuracy of the first author's name. Do not list the same citation twice.

### *PMID and DOI*

Pleased provide PubMed citation numbers to the reference list, *e.g.*, PMID and DOI, which can be found at http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed and http://www.crossref.org/SimpleTextQuery/, respectively. The numbers will be used in E-version of this journal.

### *Style for journal references*

Authors: the name of the first author should be typed in bold-faced letters. The family name of all authors should be typed with the initial letter capitalized, followed by their abbreviated first and middle initials. (For example, Lian-Sheng Ma is abbreviated as Ma LS, Bo-Rong Pan as Pan BR). The title of the cited article and italicized journal title (journal title should be in its abbreviated form as shown in PubMed), publication date, volume number (in black), start page, and end page [PMID: 11819634 DOI: 10.3748/wjg.13.5396].

### *Style for book references*

Authors: the name of the first author should be typed in bold-faced letters. The surname of all authors should be typed with the initial letter capitalized, followed by their abbreviated middle and first initials. (For example, Lian-Sheng Ma is abbreviated as Ma LS, Bo-Rong Pan as Pan BR) Book title. Publication number. Publication place: Publication press, Year: start page and end page.

### *Format*

**Journals**

*English journal article (list all authors and include the PMID where applicable)*

1 **Jung EM**, Clevert DA, Schreyer AG, Schmitt S, Rennert J, Kubale R, Feuerbach S, Jung F. Evaluation of quantitative contrast harmonic imaging to assess malignancy of liver tumors: A prospective controlled two-center study. *World J Gastroenterol* 2007; **13**: 6356-6364 [PMID: 18081224 DOI: 10.3748/wjg.13.6356]

*Chinese journal article (list all authors and include the PMID where applicable)*

2 **Lin GZ**, Wang XZ, Wang P, Lin J, Yang FD. Immunologic effect of Jianpi Yishen decoction in treatment of Pixu-diarrhoea. *Shijie Huaren Xiaohua Zazhi* 1999; **7**: 285-287

*In press*

3 **Tian D**, Araki H, Stahl E, Bergelson J, Kreitman M. Signature of balancing selection in Arabidopsis. *Proc Natl Acad Sci USA* 2006; In press

*Organization as author*

4 **Diabetes Prevention Program Research Group**. Hypertension, insulin, and proinsulin in participants with impaired glucose tolerance. *Hypertension* 2002; **40**: 679-686 [PMID: 12411462 PMCID:2516377 DOI:10.1161/01.HYP.0000035706.28494.09]

*Both personal authors and an organization as author*

5 **Vallancien G**, Emberton M, Harving N, van Moorselaar RJ; Alf-One Study Group. Sexual dysfunction in 1, 274 European men suffering from lower urinary tract symptoms. *J Urol* 2003; **169**: 2257-2261 [PMID: 12771764 DOI:10.1097/01.ju.0000067940.76090.73]

*No author given*

6 21st century heart solution may have a sting in the tail. *BMJ* 2002; **325**: 184 [PMID: 12142303 DOI:10.1136/bmj.325.7357.184]

*Volume with supplement*

7 **Geraud G**, Spierings EL, Keywood C. Tolerability and safety of frovatriptan with short- and long-term use for treatment of migraine and in comparison with sumatriptan. *Headache* 2002; **42** Suppl 2: S93-99 [PMID: 12028325 DOI:10.1046/j.1526-4610.42.s2.7.x]

*Issue with no volume*

8 **Banit DM**, Kaufer H, Hartford JM. Intraoperative frozen section analysis in revision total joint arthroplasty. *Clin Orthop Relat Res* 2002; (**401**): 230-238 [PMID: 12151900 DOI:10.1097/00003086-200208000-00026]

*No volume or issue*

9 Outreach: Bringing HIV-positive individuals into care. *HRSA Careaction* 2002; 1-6 [PMID: 12154804]

**Books**

*Personal author(s)*

10 **Sherlock S**, Dooley J. Diseases of the liver and billiary system. 9th ed. Oxford: Blackwell Sci Pub, 1993: 258-296

*Chapter in a book (list all authors)*

11 **Lam SK**. Academic investigator's perspectives of medical treatment for peptic ulcer. In: Swabb EA, Azabo S. Ulcer disease: investigation and basis for therapy. New York: Marcel Dekker, 1991: 431-450

*Author(s) and editor(s)*

12 **Breedlove GK**, Schorfheide AM. Adolescent pregnancy. 2nd ed. Wieczorek RR, editor. White Plains (NY): March of Dimes Education Services, 2001: 20-34

*Conference proceedings*

13 **Harnden P**, Joffe JK, Jones WG, editors. Germ cell tumours V. Proceedings of the 5th Germ cell tumours Conference; 2001 Sep 13-15; Leeds, UK. New York: Springer, 2002: 30-56

*Conference paper*

14 **Christensen S**, Oppacher F. An analysis of Koza's computational effort statistic for genetic programming. In: Foster JA, Lutton E, Miller J, Ryan C, Tettamanzi AG, editors. Genetic programming. EuroGP 2002: Proceedings of the 5th European Conference on Genetic Programming; 2002 Apr 3-5; Kinsdale, Ireland. Berlin: Springer, 2002: 182-191

**Electronic journal** (list all authors)

15 Morse SS. Factors in the emergence of infectious diseases. Emerg Infect Dis serial online, 1995-01-03, cited 1996-06-05; 1(1): 24 screens. Available from: URL: http://www.cdc.gov/ncidod/eid/index.htm

**Patent** (list all authors)

16 **Pagedas AC**, inventor; Ancel Surgical R&D Inc., assignee. Flexible endoscopic grasping and cutting device and positioning tool assembly. United States patent US 20020103498. 2002 Aug 1

### *Statistical data*

Write as mean ± SD or mean ± SE.

### *Statistical expression*

Express *t* test as *t* (in italics), *F* test as *F* (in italics), chi square test as $\chi^2$ (in Greek), related coefficient as *r* (in italics), degree of freedom as υ (in Greek), sample number as *n* (in italics), and probability as *P* (in italics).

### *Units*

Use SI units. For example: body mass, m (B) = 78 kg; blood pressure, p (B) = 16.2/12.3 kPa; incubation time, t (incubation) = 96 h, blood glucose concentration, c (glucose) 6.4 ± 2.1 mmol/L; blood CEA mass concentration, p (CEA) = 8.6 24.5 μg/L; $CO_2$ volume fraction, 50 mL/L $CO_2$, not 5% $CO_2$; likewise for 40 g/L formaldehyde, not 10% formalin; and mass fraction, 8 ng/g, *etc.* Arabic numerals such as 23, 243, 641 should be read 23 243 641.

The format for how to accurately write common units and quantums can be found at: http://www.wjgnet.com/2308-3840/g_info_20100725073806.htm.

***Abbreviations***

Standard abbreviations should be defined in the abstract and on first mention in the text. In general, terms should not be abbreviated unless they are used repeatedly and the abbreviation is helpful to the reader. Permissible abbreviations are listed in Units, Symbols and Abbreviations: A Guide for Biological and Medical Editors and Authors (Ed. Baron DN, 1988) published by The Royal Society of Medicine, London. Certain commonly used abbreviations, such as DNA, RNA, HIV, LD50, PCR, HBV, ECG, WBC, RBC, CT, ESR, CSF, IgG, ELISA, PBS, ATP, EDTA, mAb, can be used directly without further explanation.

***Italics***

Quantities: *t* time or temperature, *c* concentration, *A* area, *l* length, *m* mass, *V* volume.
Genotypes: *gyrA*, *arg* 1, *c myc*, *c fos*, *etc.*
Restriction enzymes: *Eco*RI, *Hin*dI, *Bam*HI, *Kbo* I, *Kpn* I, *etc.*
Biology: *H. pylori*, *E coli*, *etc.*

***Examples for paper writing***

All types of articles' writing style and requirement will be found in the link: http://www.wjgnet.com/esps/NavigationInfo.aspx?id=15

## RESUBMISSION OF THE REVISED MANUSCRIPTS

Authors must revise their manuscript carefully according to the revision policies of Baishideng Publishing Group Co., Limited. The revised version, along with the signed copyright transfer agreement, responses to the reviewers, and English language Grade A certificate (for non-native speakers of English), should be submitted to the online system *via* the link contained in the e-mail sent by the editor. If you have any questions about the revision, please send e-mail to esps@wjgnet.com.

***Language evaluation***

The language of a manuscript will be graded before it is sent for revision. (1) Grade A: priority publishing; (2) Grade B: minor language polishing; (3) Grade C: a great deal of language polishing needed; and (4) Grade D: rejected. Revised articles should reach Grade A.

***Copyright assignment form***

Please download a Copyright assignment form from http://www.wjgnet.com/2308-3840/g_info_20100725073726.htm.

***Responses to reviewers***

Please revise your article according to the comments/suggestions provided by the reviewers. The format for responses to the reviewers' comments can be found at: http://www.wjgnet.com/2308-3840/g_info_20100725073445.htm.

***Proof of financial support***

For papers supported by a foundation, authors should provide a copy of the approval document and serial number of the foundation.

***Statement about anonymous publication of the peer reviewers' comments***

In order to increase the quality of peer review, push authors to carefully revise their manuscripts based on the peer reviewers' comments, and promote academic interactions among peer reviewers, authors and readers, we decide to anonymously publish the reviewers' comments and author's responses at the same time the manuscript is published online.

## PUBLICATION FEE

*WJTM* is an international, peer-reviewed, OA online journal. Articles published by this journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium and format, provided the original work is properly cited. The use is non-commercial and is otherwise in compliance with the license. Authors of accepted articles must pay a publication fee. Publication fee: 600 USD per article. All invited articles are published free of charge.